INVESTIGACION Y CIENCIA

Edición en español de **SCIENTIFIC AMERICAN**



EXHIBICION DE CORTEJO DEL GALLO DE LAS ARTEMISAS

Julio 1978

170 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 22
Julio 1978

INVESTIGACION Y
CIENCIA
Edición en español de SCIENTIFIC AMERICAN

SCIENTIFIC AMERICAN

COMITE DE REDACCION Gerard Piel (Presidente); Dennis Flanagan, Francis Bello, Philip Morrison; Trudy E. Bell; Judith Friedman; Brian P. Hayes; Jonathan B. Piel; John Purcell; James T. Rogers; Armand Schwab, Jr.; Jonathan B. Tucker; Joseph Wisnovsky

DIRECCION EDITORIAL Dennis Flanagan
DIRECCION ARTISTICA Samuel L. Howard
PRODUCCION Richard Sasso
DIRECTOR GENERAL Donald H. Miller, Jr.

INVESTIGACION Y CIENCIA

DIRECTOR Francisco Gracia Guillén
REDACCION José María Valderas Gallardo (Redactor Jefe)
M.ª Rosa García Solé (Secretaria de Redacción)
PRODUCCION César Redondo Zayas
PROMOCION Y PUBLICIDAD Pedro Clotas Cierco

EDITA Prensa Científica, S.A.
Calabria, 235-239
Barcelona-29
ESPAÑA

Colaboradores de este número:

Asesoramiento y traducción:

Josep Rucabado: *Armas de radiación intensiva;* Miquel Gich: *Cuando el mar Negro se desecaba;* Manuel Puigcerver: *La radiación cósmica de fondo y el nuevo movimiento del éter;* Joandomènec Ros: *El sistema de apareamiento de lek del gallo de las artemisas;* José Manuel García de la Mora: *El principio de adyacencia en la percepción visual;* Monique Robert: *Uniones intercelulares;* Antonio Blanco: *Tecnología hidráulica romana;* Luis Bou; *Juegos matemáticos;* José E. Myro y Alicia Menéndez: *Taller y laboratorio.*

Ciencia y sociedad:
C. Rodríguez Barrueco.

Libros:
Enrique Gadea y Philip Morrison.



LA PORTADA

La ilustración de la cubierta muestra a un macho de gallo de las artemisas exhibiéndose en su territorio de reproducción en Wyoming, a principios de primavera. El macho se encuentra en el estadio llamado pavoneo de una exhibición repetitiva de tres segundos; su saco esofágico hinchado, que se halla bajo los dos abultamientos prominentes de piel desnuda, está a punto de contraerse y de producir un sonido similar al chasquido que se origina al descorchar una botella. Cuando el macho de gallo de las artemisas se pavonea en presencia de posibles compañeras, puede exhibirse más de 30 veces durante seis minutos (véase "El sistema de apareamiento de lek del gallo de las artemisas", de R. Haven Wiley, Jr., en este mismo número).

Suscripciones:

Prensa Científica, S.A.
Departamento de suscripciones
Calabria 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 41

Distribución para España:

Distribuciones de Enlace, S.A.
Ausias March, 49
Barcelona-10 (España)

Distribución para los restantes países:

Editorial Labor, S.A.
Calabria, 235-239
Barcelona-29 (España)

Precio de venta de este número:
España: 170 pesetas
Extranjero: 2 U.S. $

Condiciones de suscripción:
España:
Un año (12 números): 1870 pesetas
Extranjero:
Un año (12 números): 30 U.S. $

Ejemplar atrasado ordinario:
190 pesetas

Ejemplar atrasado extraordinario:
270 pesetas.

Difusión controlada por **OJD**

PROCEDENCIA DE LAS ILUSTRACIONES

Dibujo de la portada de Enid Kotschnig

Página	Fuente	Página	Fuente
5-8	Ilil Arbel	76	Barbara E. Hull
9	Sovfoto	78	Patricia J. Wynne
10	Ilil Arbel	79	Barbara E. Hull *(arriba),* L. Andrew Staehelin *(centro),* Daniel S. Fripnd *(abajo)*
14	National Aeronautics and Space Administration	80	Patricia J. Wynne
16-20	Bunji Tagawa	81	Lelio Orci
21	Kenneth J. Hsü	82	Barbara E. Hull *(arriba),* Patricia J. Wynne *(abajo)*
22-23	Bunji Tagawa	83	Camillo Peracchia *(arriba);* Barbara E. Hull *(centro); The Journal of Cell Biology,* Rockefeller University Press *(abajo)*
24	Kenneth J. Hsü *(arriba),*	84-86	Patricia J. Wynne
25-26	Bunji Tagawa *(abajo)*	89	Norman Smith
28	Lawrence Berkeley Laboratory, Universidad de California	90-92	Tom Prentiss
29-38	Gabor Kiss	93	Norman Smijh
45-49	L. Cornudella y A. Sellés	94-95	Tom Prentiss
50	Collier's	97-100	Jerome Kuhl
51	L. Cornudella	102	Design Photographers International, Inc.
52	L. Cornudella y A. Sellés	103-107	Michael Goodman
55	R. Haven Wiley, Jr.		
56-59	Alan D. Iselin		
60-61	R. Haven Wiley, Jr.		
62-63	Alan D. Iselin		
67	Jon Brenneis		
68-74	Ilil Arbel		

Dep. legal: B. 38.999-76
Fotocomposición: Fort, S.A.
Rosellón, 33 - Barcelona-29
Cayfosa. Santa Perpetua de Moguda. Barcelona.

Printed in Spain - Impreso en España

Los autores

FRED M. KAPLAN ("Armas de radiación intensiva") participa en el proyecto sobre control de armamento en el Centro de Estudios Internacionales del Instituto de Tecnología de Massachusetts. Estudió en Oberlin College, en donde se licenció con las máximas calificaciones en 1976. Durante los veranos de 1975 a 1977 investigó en cuestiones de defensa en la Carnegie Endowment for International Peace, en el Military Audit Project y en el Instituto de Estudios Políticos norteamericano. Kaplan ha escrito artículos y colaboraciones sobre política estadounidense y rusa de defensa en varias publicaciones periódicas.

KENNETH J. HSU ("Cuando el mar Negro se desecaba") dirige el Instituto Geológico del Politénico federal de Zurich. Nació en Nankin y realizó sus estudios en la Universidad central de China. En 1948 se trasladó a Estados Unidos para continuar sus estudios de geología en la Universidad estatal de Ohio y en la Universidad de California en Los Angeles. Después de haber trabajado para la Shell y haber enseñado en la Universidad estatal de Nueva York y en la de California en Riverside, pasó al Politécnico suizo. en 1967. Desde entonces se ha dedicado primordialmente al estudio de los sedimentos marinos. Hsü desempeñó el cargo de codirector científico en dos expediciones al Mediterráneo con el Deep Sea Drilling Project y, en 1975, se ocupó de los problemas sedimentológicos del *Glomar Challenger* en una expedición al mar Negro.

RICHARD A. MULLER ("La radiación cósmica de fondo y el nuevo movimiento del éter") es físico investigador asociado en el Laboratorio de Ciencias del Espacio de la Universidad de California en Berkeley y en el Laboratorio Lawrence de la misma. Estudió en las Universidades de Columbia y Berkeley, donde obtuvo su doctorado en Física de partículas elementales en 1969. Desde entonces, sus variados intereses en investigación han incluido (aparte de las mediciones de la radiación cósmica de fondo descritas en este artículo) la puesta a punto de un espejo de telescopio que automáticamente compensa la distorsión atmosférica, la búsqueda de quarks con carga unidad y el desarrollo de una nueva técnica de datación mediante isótopos radiactivos.

L. CORNUDELLA ("El nucleosoma") es investigador en el departamento de química macromolecular del Consejo Superior de Investigaciones Científicas (CSIC) en Barcelona. Cursó la licenciatura de ciencias químicas en la facultad de ciencias de la Universidad de Barcelona, donde se especializó en bioquímica y obtuvo el grado de doctor en 1969. Un año más tarde se desplazó a los Estados Unidos mediante una beca postdoctoral concedida por la fundación Juan March, para trabajar en el laboratorio de biología celular del Kimball Research Institute de Nueva York bajo la dirección del profesor A. O. Pogo. Durante su permanencia en dicho laboratorio, su labor investigadora se dirigió hacia el estudio de la biogénesis de las moléculas informacionales y de los mecanismos que regulan el transporte de la información genética en células eucariotas. De regreso a España en 1973 ingresó en el CSIC. Su centro de interés lo constituye el estudio de la organización estructural de la cromatina en organismos superiores y el aislamiento y caracterización de mensajes genéticos específicos en células germinales en proceso de diferenciación.

R. HAVEN WILEY, JR. ("El sistema de apareamiento de lek del gallo de las artemisas") es profesor asociado de zoología en la Universidad de Carolina del Norte en Chapel Hill. Estudió en el Harvard College y obtuvo su doctorado en comportamiento animal por la Universidad Rockefeller en 1970. Después de un año de postdoctorado en Rockefeller se incorporó a la facultad en Chapel Hill, donde dirige hoy la Behavioral Research Station del Jardín Botánico de Carolina del Norte. Los campos de investigación de Wiley, además del sistema de apareamiento de lek en el gallo de las artemisas, son las adaptaciones ecológicas de la estructura social de estorninos y chochines, los sistemas de comunicaciones acústicas de las aves y el comportamiento que regula el control de las relaciones espaciales en las sociedades animales.

WALTER C. GOGEL ("El principio de adyacencia en la percepción visual") es profesor de psicología en la Universidad de California en Santa Bárbara. Estudió física en el Marietta College y se especializó en psicología experimental, en la Universidad de Chicago, bajo la dirección de Eckhard H. Hess. Doctorado en 1951, dedicó los 14 años siguientes a investigar sobre la percepción espacial en el U.S. Army Medical Research Laboratory de Fort Knox y en el Civil Aeromedical Research Institute de Oklahoma City. Entró a formar parte del claustro universitario de Santa Bárbara en 1965.

L. ANDREW STAEHELIN y BARBARA E. HULL ("Uniones intercelulares") son ambos biólogos celulares interesados en la estructura y función de las membranas biológicas. Staehelin es profesor adjunto en la Universidad de Colorado en Boulder. Nació en Australia, realizó sus estudios secundarios en Suiza y se doctoró en el politécnico helvético en 1966. Durante los tres años siguientes, estuvo trabajando en el Departamento de Investigación Científica y Técnica de Nueva Zelanda en torno a la estructura de la pared celular vegetal, trasladándose a los Estados Unidos en 1969 para ocupar una plaza de becario de investigación en el laboratorio de Keith R. Porter, en Harvard. Al año siguiente, se pasó, junto con Porter, al recientemente creado departamente de biología molecular, celular y del desarrollo de la Universidad de Colorado. Desde entonces, las investigaciones de Staehelin se han centrado sobre las uniones que coordinan las actividades celulares de los tejidos y sobre la estructura de las membranas fotosintéticas. Staehelin se halla actualmente disfrutando de un año sabático de licencia en la universidad de Friburgo en Alemania Federal. Hull es adjunto de investigación en la sección de biología celular de la Facultad de Medicina de la Universidad de Yale. Cursó sus estudios en el Smith College y posteriormente en la Universidad de Colorado, donde realizó trabajos de investigación en el laboratorio de Staehelin, doctorándose en 1976.

NORMAN SMITH ("Tecnología hidráulica romana") explica historia de la tecnología en el Imperial College of Science and Technology, de Londres. Nacido en Southampton, estudió en la Universidad de Bristol, donde recibió el grado de doctor en ingeniería civil, en 1962. Después enseñó ingeniería en la Universidad de Canterbury, en Nueva Zelanda; pasó al Imperial College en 1965. Su reciente labor de investigación se ha centrado en la historia de la construcción de presas y evolución de la tecnología hidráulica, tema sobre el que ha escrito tres libros. Cuando no se ocupa de escribir o de enseñar, combina el trabajo de campo histórico con los viajes y la fotografía.

Armas de radiación intensiva

La nueva arma nuclear táctica norteamericana, ampliamente conocida por "bomba de neutrones", tiene una utilidad militar dudosa. Y, lo que es más importante, su uso en una guerra en Europa conduciría a una guerra completamente nuclear

Fred M. Kaplan

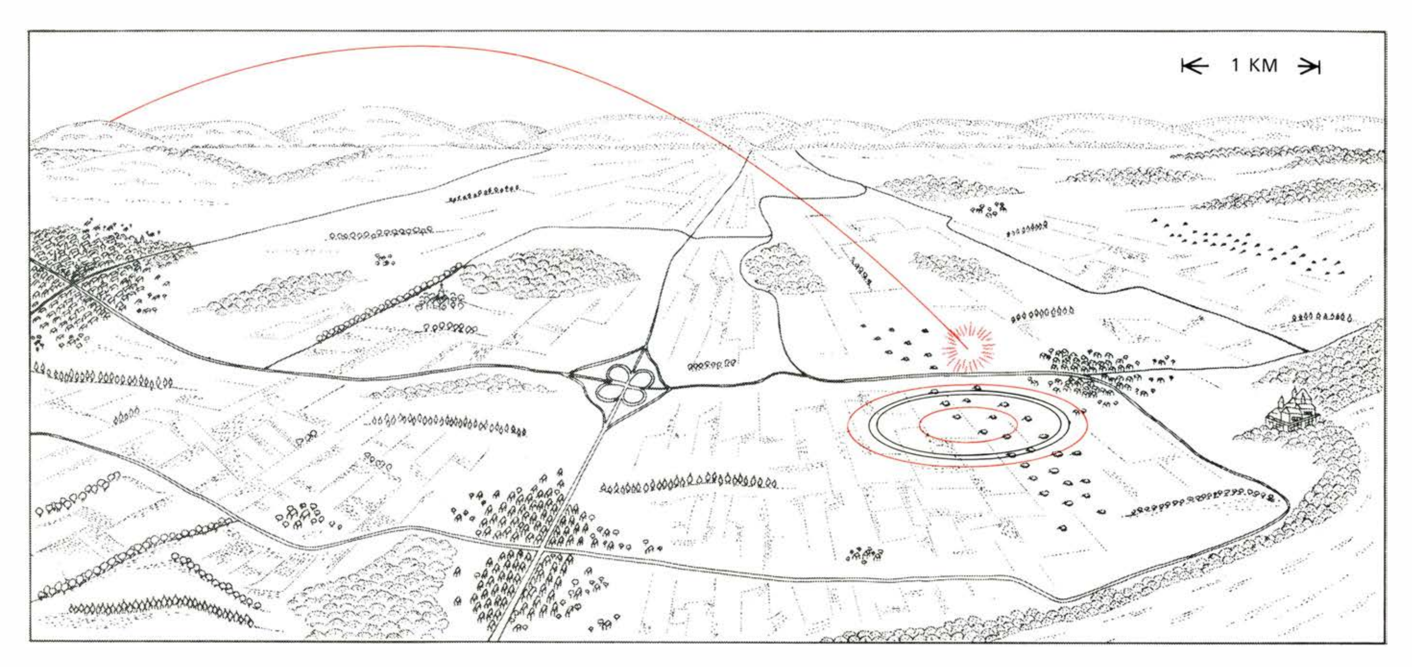
La bomba de radiación intensiva (o, como amplia y un tanto erróneamente se la conoce, la bomba de neutrones) constituye el más reciente desarrollo de la investigación militar norteamericana para disponer de un arma nuclear más utilizable y más "limpia". Este nuevo tipo de bomba, algunas de cuyas versiones podrían estar terminadas en 1979, se ha diseñado para matar más tropas enemigas por kilotón de potencia explosiva detonada sobre el campo de batalla que los tipos de armas nucleares corrientemente desarrollados para tal propósito, minimizando simultáneamente los daños colaterales, o no intencionados, a los edificios, al territorio, a las tropas aliadas y a los civiles que se encuentren en los alrededores.

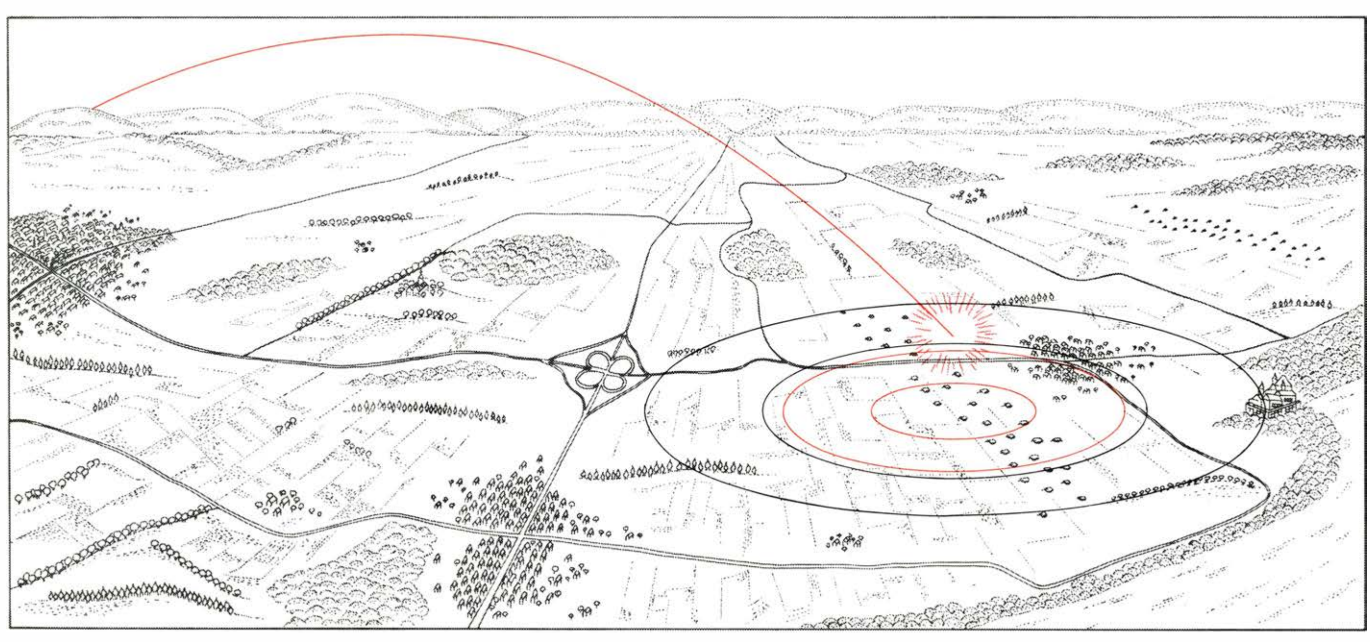
Muchos militares sostienen que mediante el uso de tales bombas de radiación intensiva, más precisas y refinadas, la "guerra nuclear limitada" podría quedar efectivamente limitada, y su efecto, confinado prácticamente al campo de batalla. Al igual que la generación de armas nucleares tácticas que la precedió, piensa utilizarse la bomba de radiación intensiva en una posible guerra en territorio europeo entre las naciones de la Organización del Tratado del Atlántico Norte (OTAN), incluyendo a los Estados Unidos, y las naciones del Pacto de Varsovia, comprendida la Unión Soviética. El presupuesto militar de la Administración Carter para el año fiscal 1979 destina un monto de gastos sin precedentes a las fuerzas norteamericanas preparadas para una eventual guerra europea. Por tanto, la nueva arma nuclear táctica merece un debate detallado, habida cuenta de la extraordinaria publicidad concedida al arma y el desconcierto generalizado en torno a dicho tema. ¿Cómo se concibió el arma? ¿Cómo funciona? ¿Cuáles son sus efectos? ¿Cuál su utilidad militar? ¿Debe producirse y desarrollarse?

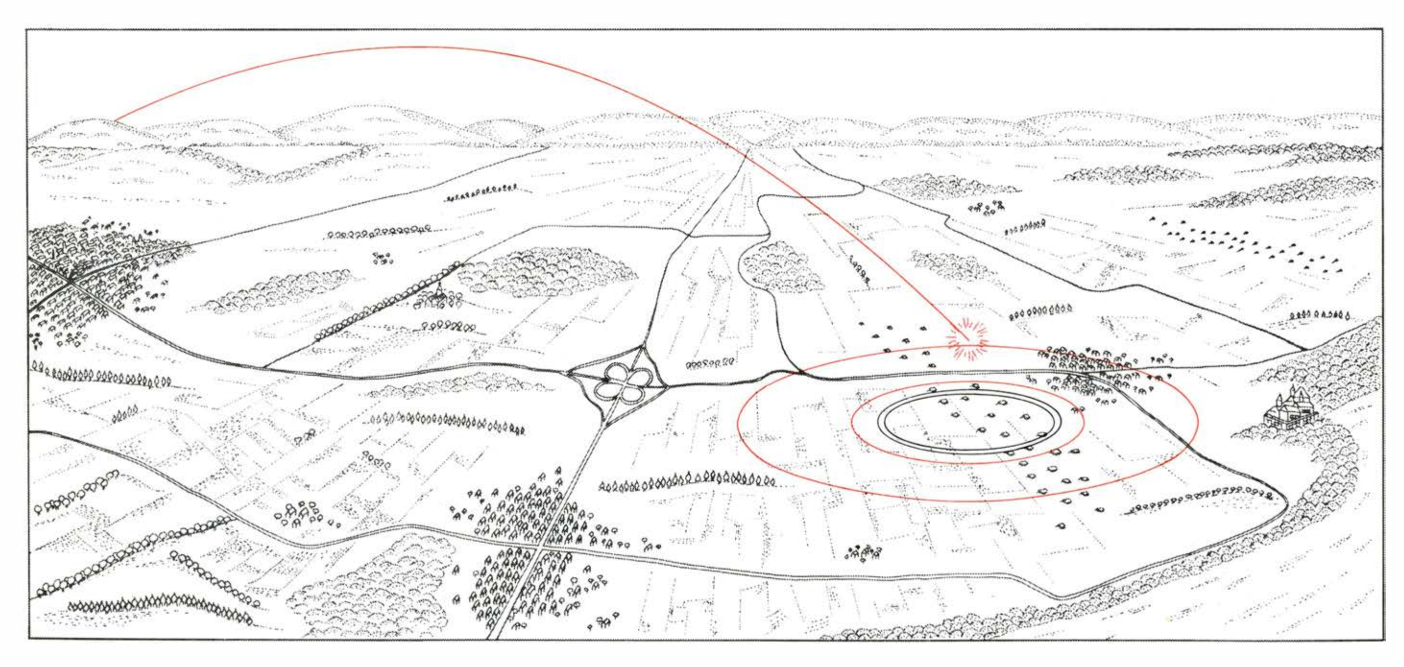
Importa destacar, desde un principio, que la idea de una bomba de neutrones no es nada nuevo. La posibilidad de desarrollar un arma nuclear táctica de este tipo fue conocida poco después de la invención de la bomba de hidrógeno o de fusión, en las postrimerías de la década 1940-1949. Varios científicos, dedicados a la investigación de armas nucleares, principalmente en el Lawrence Livermore Laboratory, desarrollaron la noción de un arma de radiación intensiva a lo largo de las dos décadas siguientes, y haciendo fuerza con otras personas presionaron políticamente para el posterior desarrollo y despliegue del artefacto.

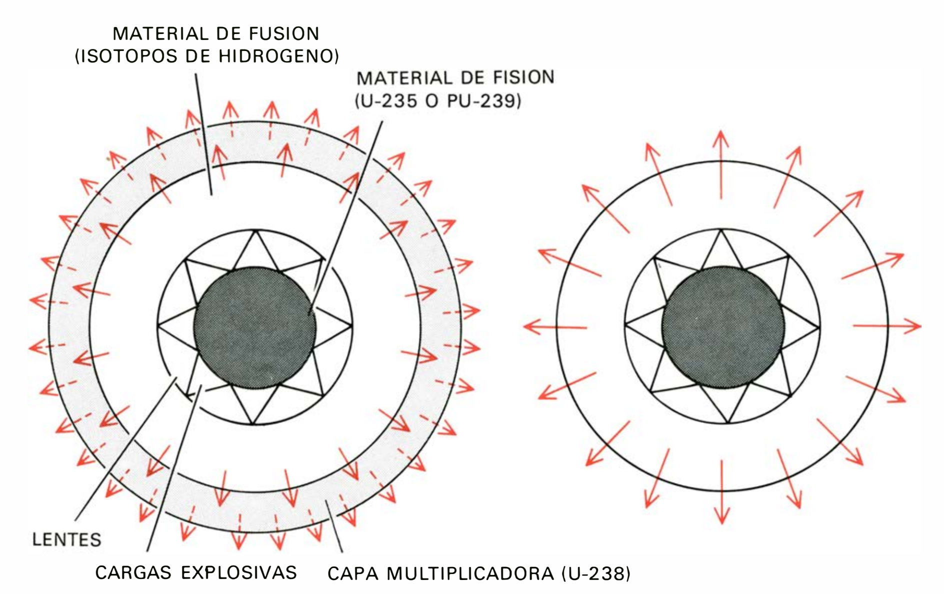
Sin embargo, hasta principios de los años 60 el Secretario de Defensa Robert S. McNamara no ordenó un estudio global sobre las perspectivas de las armas nucleares tácticas. En base a este estudio, y a varias maniobras de simulación de guerra, llegó a la conclusión de que una guerra nuclear en el escenario europeo constituiría una batalla perdida por ambos lados. Millones de civiles morirían, y el uso de tales armas no significaría necesariamente una superioridad de la OTAN en una guerra europea. Las armas nucleares tácticas, lejos de sustituir a los combatientes y a la potencia de fuego convencional, requerían mayores contingentes de tropa, de forma que los soldados de la OTAN muertos por la respuesta nuclear de la Unión Soviética fueran rápidamente reemplazados. Se dedujo finalmente que si las fuerzas del Pacto de Varsovia movilizaran las tropas en el frente según el estilo escalón que tienen previsto y si las fuerzas de la OTAN

ALGUNOS EFECTOS de tres tipos diferentes de bombas nucleares tácticas quedan de manifiesto en las escenas convencionales de la página opuesta. En los tres casos, la acción se desarrolla en una zona semirrural de Alemania Occidental, en donde los oficiales de la Organización del Tratado del Atlántico Norte prevén una invasión por las fuerzas de las naciones del Pacto de Varsovia. Cada saliente de esta ofensiva estaría conducido presumiblemente por una fuerza de ataque fomada por centenares de tanques rusos, que se supone se desplazarían en dos o tres escalones, en el primero de los cuales los tanques estarían separados por una distancia de 75 a 100 metros, aproximadamente. Justo a la derecha de la parte central de cada escena figura un escalón de este tipo; a tres kilómetros detrás del primero, en la parte superior derecha, sigue un segundo escalón. En los tres casos, fuerzas de la OTAN situadas a unos 130 kilómetros han lanzado un misil Lance provisto de una cabeza nuclear táctica contra el primer escalón de tanques rusos. La escena superior muestra los efectos explosivos y térmicos estimados *(en negro)* y los efectos de las radiaciones iniciales *(en color)* de una bomba nuclear de fisión, de una potencia explosiva de un kilotón, detonada sobre el objetivo a una altura de 500 metros. La ilustración central y la inferior muestran, respectivamente, los efectos correspondientes de una bomba de fisión de 10 kilotones y de una bomba de fisión-fusión de radiación intensiva (la bomba de neutrones) de un kilotón. En los tres casos, el círculo negro interior delimita la zona dañada por los efectos explosivos, en la cual la onda de choque alcanza una sobrepresión de 350 gramos por centímetro cuadrado, suficiente para destruir la mayoría de los edificios. El círculo negro exterior corresponde al área en donde la radiación térmica es suficiente para causar quemaduras de segundo grado a las personas que no lleven protección externa. El círculo interior de color indica el límite de la zona expuesta a un mínimo de 8000 rads de radiación inicial (principalmente neutrones) suficiente para causar una "incapacitación permanente inmediata" a los soldados o, en el caso de suponer una protección corriente en los tanques, una "incapacitación temporal inmediata" a los conductores. El círculo exterior de color corresponde a una dosis de radiación inicial de 150 rads, suficiente para matar alrededor del 10 por ciento de las personas expuestas y provocar una alta incidencia de cáncer en los sobrevivientes. La escala se indica en la parte superior derecha de la primera figura. En esta representación gráfica esquematizada no se han considerado los efectos de la radiación residual (lluvia radiactiva).


1 KM







DOS TIPOS DE BOMBA FISION-FUSION, confrontados esquemáticamente. En ambos casos, un conjunto de cargas explosivas químicas constituye el detonante de la bomba, provocando las reacciones de fisión en su núcleo, las cuales a su vez constituyen el detonante para las reacciones de fusión que tienen lugar en una capa envolvente. El dispositivo normal de fisión-fusión *(izquierda)* está recubierto por una capa adicional de uranio 238, isótopo no fisionable, que aumenta considerablemente la potencia explosiva de la bomba al capturar la mayoría de los neutrones rápidos *(flechas continuas en color)* emitidos en el proceso de fusión. Los núcleos de U-238 se fisionan luego emitiendo gran cantidad de neutrones "térmicos" *(flechas discontinuas en color)*. La mayoría de las armas nucleares estratégicas se basan en este concepto. La bomba de radiación intensiva *(derecha)* elimina la capa de U-238. De este modo, está bomba libera una mayor proporción de neutrones rápidos a expensas de su potencia explosiva. Los neutrones lentos quedan capturados por los núcleos atómicos en la atmósfera, mucho más rápidamente que los neutrones rápidos. La bomba de radiación intensiva produce muchos más neutrones que la bomba de fisión-fusión normal.

continuaran con sustituciones individuales dentro de la estructura actual de divisiones, una guerra nuclear en Europa probablemente favorecería a la Unión Soviética y a sus aliados, incluso en el caso de que la OTAN poseyera más y "mejores" armas nucleares.

Además, se consideró muy elevado el riesgo de escalada hacia una guerra nuclear estratégica completa, entre los Estados Unidos y la Unión Soviética, como consecuencia de esta estrategia; y ello básicamente por dos razones. En primer lugar, porque la frontera entre guerra nuclear y guerra convencional estaba clara en aquel momento; el obstaculizar la distinción entre guerra nuclear táctica y estratégica hubiera creado una ambigüedad considerable, pudiendo conducir a recelos mutuos, a tensiones y, posiblemente, a impactos nucleares estratégicos de carácter preventivo. En segundo lugar, la Unión Soviética tenía muchos misiles balísticos de alcance intermedio (IRBM) con cabezas nucleares emplazados en su territorio, algunos de los cuales en las mismas zonas que las ocupadas por los misiles balísticos intercontinentales (ICBM); en los primeros estadios de una hipotética guerra nuclear europea, la OTAN se vería fuertemente inducida a destruir, en prevención, los IRBM instalados en territorio soviético, desencadenando posiblemente un enfrentamiento nuclear estratégico entre las dos superpotencias.

Después de ponderar estas consideraciones, McNamara se inclinó por una política de aumento de las capacidades de combate convencional, o no nuclear, y cortó el suministro de fondos para una nueva generación de armas nucleares tácticas. (No obstante, aceptó el misil Lance con cabeza nuclear, a causa de su mayor alcance y su consiguiente menor vulnerabilidad.) Durante el tiempo en que Melvin R. Laird ocupó la Secretaría de Defensa se destinó más dinero al desarrollo de una nueva generación de armas nucleares tácticas, pero esencialmente prevaleció la actitud negativa hacia la modernización de estas armas hasta que James R. Schlesinger fue nombrado Secretario en 1973.

Schlesinger, durante el desempeño de su cargo anterior como Presidente de la Comisión de Energía Atómica, mostró un entusiasmo considerable hacia las armas nucleares tácticas. Sin embargo, al llegar a la Secretaría de Defensa, su interés parecía haber disminuido. El mismo se sintió obligado, al parecer, a hacer algunas concesiones a los defensores de las armas nucleares tácticas, que formaban un conjunto de intereses convergentes constituido por la Comisión de Energía Atómica, el Comité del Congreso sobre Energía Atómica, los laboratorios de armamento, determinados departamentos militares y la División de Energía Atómica de la Secretaría de Defensa. Schlesinger, que tenía poco poder de negociación en la Casa Blanca durante las Administraciones Nixon y Ford, tuvo que dirigir su propia y compleja coalición. Para ganar el apoyo de estos grupos desiguales hacia sus planes de incrementar las fuerzas convencionales de la OTAN, les proporcionó fondos destinados a la modernización de las armas nucleares tácticas.

La coalición en favor de un programa de modernización de las armas nucleares tácticas fue activamente estimulada por el propio empeño de Schlesinger en aumentar el abanico de "opciones" de los Estados Unidos en la planificación de la fuerza nuclear. Gracias en parte a las nuevas tecnologías, tales como los sistemas de guía inercial de gran precisión para misiles, Schlesinger programó de nuevo las armas nucleares estratégicas a fin de dotarlas de mayor "flexibilidad" y de mayor capacidad de "impactos selectivos", creando nuevos "paquetes de objetivos" que fueran bastante más diversificados que los disponibles por los planificadores de defensa de la década anterior. Junto con la expansión de las opciones estratégicas, ordenó un aumento de las opciones disponibles de lucha en una guerra nuclear en Europa. Este nuevo refuerzo apareció como el apoyo oficial a los militares que habían empezado a pensar seriamente en la posibilidad de combatir y ganar una guerra nuclear limitada y en la necesidad, bajo tales circunstancias, de limitar los daños colaterales.

Entretanto, el desarrollo del sistema de misiles antibalísticos Sprint (ABM) en Los Alamos Scientific Laboratory, a mediados de la década 1960-69, y la subsiguiente declaración sobre la producción futura de ABM impuesta por el tratado SALT I de 1972, hizo pensar a algunos científicos del laboratorio de armamento en reducir la potencia de las cabezas nucleares Sprint para adaptarlas a las armas nucleares tácticas. (Los misiles Sprint antimisiles, de corto alcance, dotados de carga nuclear, están diseñados para explosionar en la atmósfera y su eficacia depende más de los neutrones liberados que de los rayos X.) Todos estos intereses diversos –de los servicios militares, de los laboratorios

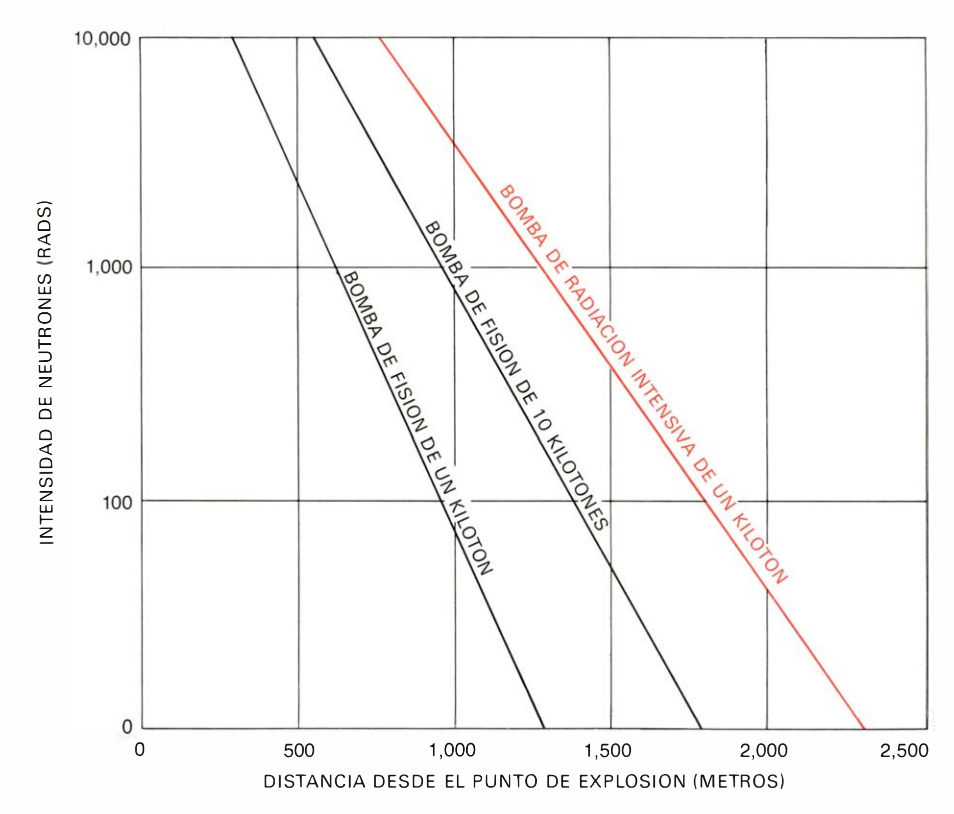
de armamento, de los comités del Congreso y del Departamento de Defensa- convergieron hasta configurar la presente situación.

Las actuales cabezas nucleares de radiación intensiva se han desarrollado para el misil Lance y para el proyectil de artillería de ocho pulgadas (203,20 milímetros). También está en proyecto una cabeza nuclear de radiación intensiva para el proyectil de artillería de 155 milímetros, aunque todavía se encuentra en las primeras etapas de desarrollo. (Una de estas bombas, por lo menos, probablemente la destinada al Lance, ha sido ya experimentada en una prueba subterránea cerca de Las Vegas.) Las cabezas nucleares corrientemente emplazadas para el Lance tienen un rango de potencia explosiva que abarca desde un kilotón hasta 100 kilotones; las cargas de artillería nuclear de ocho pulgadas varían desde 5 hasta 10 kilotones. La nueva versión en radiación intensiva para el Lance tendrá dos potencias, que podrán prefijarse simplemente apretando unos botones: una potencia considerablemente inferior a un kilotón y la otra ligeramente superior a un kilotón. El proyectil de artillería nuclear de ocho pulgadas tendrá tres potencias, que variarán desde una muy inferior a un kilotón hasta aproximadamente dos kilotones.

Los efectos de una explosión nuclear consisten en el daño producido por la mera explosión (una onda de choque de sobrepresión), por la radiación térmica (calor), por la radiación instantánea (la mayor parte neutrones y rayos gamma) y la radiación residual (la radiactividad que procede de la lluvia de productos de fisión). La energía liberada por una explosión por fisión se divide en varias fracciones; en números redondos: 50 por ciento en onda de choque, 35 por ciento en radiación térmica, 5 por ciento en radiación instantánea y 10 por ciento en radiación residual. En una hipotética bomba de fusión pura, los efectos serían: 20 por ciento en onda de choque o radiación térmica, 80 por ciento en radiación instantánea (la mayor parte neutrones) y relativamente poca radiación residual (la cantidad exacta depende de las características del suelo que alcanza la explosión). La reacción de fusión que tiene lugar entre los iones de deuterio y de tritio (dos isótopos pesados del hidrógeno) viene acompañada por la liberación de neutrones rápidos o de muy alta energía. La energía de estos neutrones es de unos 14 millones de electrón voltios (MeV), muy superior a los neutrones ya bastante rápidos de 2 MeV liberados en una reacción de fisión típica. Los neutrones se ven enlentecidos y, con el tiempo, capturados por los residuos de la propia bomba, por otros objetos que encuentra en su recorrido y por el aire. Cuanto más rápidos sean, tantas más colisiones sufrirán antes de su total captura. Además, la fusión produce 10 veces más neutrones por kilotón de potencia explosiva que la fisión. Así, los neutrones producidos por una bomba de fusión tienen mayor intensidad de radiación, y alcanzan mayores distancias antes de ser absorbidos, que los producidos por una bomba de fisión.

La actual bomba de radiación intensiva es una bomba de fisión-fusión. La proporción fisión-fusión no es exactamente la misma entre la cabeza para el Lance y la cabeza para la artillería de ocho pulgadas, pero el proceso de detonación es el mismo en ambos artefactos. En el proceso de detonación del arma, la reacción de fisión constituye el detonante para la reacción de fusión, la cual libera muchos neutrones rápidos. Esta es la razón por la que suele denominarse bomba de neutrones a la bomba de radiación intensiva. Tal apelativo es correcto, toda vez que la bomba de radiación intensiva libera muchos más neutrones que otras bombas de potencia equivalente. Pero se olvida que, en la detonación de esta bomba, se libera también gran cantidad de energía bajo otras formas. (De hecho, toda bomba nuclear inferior a dos kilotones puede llamarse bomba de neutrones, en el sentido de que en la zona correspondiente al radio letal de la bomba, aun cuando se tratara de una bomba de fisión pura, la energía liberada en forma de radiación instantánea es mayor que la radiación térmica y la onda de choque, y que la radiación instantánea en forma de neutrones es superior a la radiación instantánea en forma de rayos gamma. Si tal bomba explotara en el aire a una altura de varios centenares de metros, sus efectos explosivos y térmicos en el suelo serían ligeros, aunque el daño causado por los neutrones seguiría siendo muy importante.)

La bomba de radiación intensiva está lejos de ser una bomba de fusión pura. Si nos atenemos a su potencia explosiva, las bombas de radiación intensiva de hasta un kilotón para la artillería de ocho pulgadas serían de un 50 por ciento de fisión y otro 50 por ciento de fusión. La versión de radiación intensiva para el



NEUTRONES EMITIDOS por tres tipos de armas nucleares tácticas representados en función de la distancia al punto en el que se ha producido la explosión. Las dos curvas en negro dan la intensidad de neutrones para una bomba de fisión de un kilotón y otra de 10 kilotones. La curva en color representa la intensidad para una bomba de radiación intensiva de fisión-fusión (bomba de neutrones) de un kilotón.

Lance es, aproximadamente, de 60 por ciento de fusión y un 40 por ciento de fisión. La bomba de radiación intensiva de 2 kilotones para la artillería de ocho pulgadas es de 70 a 75 por ciento de fusión. La energía liberada por las bombas de radiación intensiva para el Lance y para la artillería de ocho pulgadas de menor protencia se distribuye, en números redondos, de la siguiente forma: 40 por ciento en onda de choque, 25 por ciento en radiación térmica, 30 por ciento en radiación instantánea y 5 por ciento en lluvia radiactiva. Las bombas de radiación intensiva para la artillería de ocho pulgadas de mayor potencia producen cerca del 10 por ciento más de radiación instantánea y ligeramente menos en onda de choque, radiación térmica y radiación residual. Con otras palabras, la bomba de radiación intensiva no promete ser ni el arma libre de efectos colaterales que sus entusiastas afirman, ni el "arma capitalista final" (que destruye sólo a las personas pero no la propiedad), como teme mucha gente de los grupos para la paz.

La distinción fundamental entre la bomba de radiación intensiva y las otras armas nucleares de baja potencia dotadas de una proporción de fisión mayor estriba en que la primera libera muchos más neutrones y a una velocidad mucho mayor. La energía liberada por las bombas de radiación intensiva para el Lance y para la artillería de ocho pulgadas de poca potencia, en forma de radiación instantánea, es seis veces superior a la de las bombas de fisión de potencia equivalente, llegando a 10 en las bombas de radiación intensiva de mayor potencia para la artillería de ocho pulgadas.

Existe otra distinción, aparte del número de kilotones, entre las bombas de radiación intensiva y las de fisión-fusión de mayor potencia. Las de fisión-fusión normales (entre las que se cuentan la mayoría de las bombas nucleares estratégicas) están recubiertas por una "capa" de uranio 238 que incrementa la potencia explosiva del arma: la capa captura o frena considerablemente los neutrones rápidos liberados por el proceso de fusión. Puesto que las bombas de radiación intensiva pretenden conseguir potencias térmicas muy bajas y la liberación de muchos neutrones rápidos, no llevan la capa de U-238.

Entonces, ¿cuál se supone que debe ser la misión militar de las bombas de radiación intensiva? La mayor preocupación de muchos militares de la OTAN es la posibilidad de una invasión relámpago de las tropas del Pacto de Varsovia, dirigidas por Rusia, a través de las llanuras septentrionales de Alemania Occidental. La doctrina militar rusa y el despliegue de las fuerzas rusas sugieren que el lanzamiento de un ataque de este tipo involucraría el despliegue de miles de tanques como principio motor de la ofensiva. Para ciertos especialistas de estado mayor, un ataque de este tipo, especialmente si se realizara la movilización en un corto período de tiempo, no podría ser detenido por la OTAN sin el recurso a las armas nucleares. (Esta hipótesis ha sido acaloradamente discutida.) Durante varios años algunos oficiales de los Estados Unidos han criticado la "inviabilidad" de la mayoría de las armas nucleares estacionadas en los arsenales de Europa Occidental, poniendo particular atención en sus potencias relativamente elevadas, algunas de ellas muy superiores a los 20 kilotones que destruyeron gran parte de la ciudad de Nagasaki al final de la Segunda Guerra Mundial. Tales armas de alta potencia serían efectivas para detener a los tanques rusos, pero también podrían matar o herir gravemente a muchos soldados de la OTAN, a muchos civiles alemanes, además de devastar extensos territorios de Alemania Occidental. Por otro lado, los efectos de la radiación inducida y residual convertirían a la ocupación y recuperación del territorio afectado en un proyecto letal durante algún tiempo.

Con la bomba de radiación intensiva los militares encontraron una táctica diferente: matar a las tropas del Pacto de Varsovia dentro de los tanques, sin destruir las máquinas. Lo cual puede conseguirse, dicen, por el gran flujo de neutrones generado por las bombas de radiación intensiva.

Las dosis de radiación se miden en rads; un rad equivale a la dosis absorbida de radiación nuclear que acompaña a la liberación de 100 ergios de energía por gramo de material irradiado. Para

	NOMBRE DEL SISTEMA PROPULSOR	NUMERO DE CABEZAS EN EUROPA	POTENCIA EXPLOSIVA (KILOTONES)	ALCANCE MAXIMO (KM.)
OTAN	MISIL HONEST JOHN	196	20	40
	MISIL PERSHING	180	60-400	735
	MISIL LANCE	80	1-100	110
	MISIL SERGEANT	56	BAJA	135
	MISIL PLUTON	24	15-25	120
	MISIL SSBS-2	18	150	3000
	OBUS DE OCHO PULGADAS M-110	360	5-10	15
	OBUS DE OCHO PULGADAS M-115	27	5-10	15
	OBUS DE 155 MILIMETROS M-109	691	BAJA	15
	MISIL LANCE CON BRI	?	APROX. 1	135
	OBUS DE OCHO PULGADAS CON BRI	?	1-2	>15
	OBUS DE 155 MILIMETROS CON BRI	?	?	?
PACTO DE VARSOVIA	MISIL SS-4 SANDAL	500	1000	1900
	MISIL SS-5 SKEAN	100	1000	3700
	MISIL SS-20	20 (×3)	?	4800
	MISIL SS-1b SCUD A	880 (SS-1b, SS-1c, SS-12)	?	80
	MISIL SS-1c SCUD B		?	290
	MISIL SS-12 SCALEBOARD		>1100	800
	MISIL FROG 3-7	650	?	70
	OBUS DE 203 MILIMETROS M-55	?	?	30

ARMAS NUCLEARES TACTICAS corrientemente desplegadas en Europa por los países de la OTAN (incluidos los Estados Unidos), y por los países del Pacto de Varsovia (comprendida la Unión Soviética); la comparación alcanza sólo a la tipografía en negro. En color se relacionan las nuevas armas de radiación intensiva propuestas por los Estados Unidos para el arsenal de la OTAN. La versión para radiación intensiva del misil Lance podría tener dos potencias, una algo inferior a un kilotón, y la otra ligeramente superior. La versión para radiación intensiva de la bomba para artillería de ocho pulgadas podría tener tres potencias explosivas, variando desde bastante menos de un kilotón hasta cerca de dos kilotones. El misil balístico de alcance intermedio (IRBM) ruso lleva tres cabezas nucleares que pueden dirigirse independientemente hacia sendos objetivos.



TANQUE RUSO DE COMBATE ABIERTO designado por T-62, fotografiado en una parada militar en la Plaza Roja de Moscú. Las fuerzas del Pacto de Varsovia tienen 20.000 carros pesados (la mayoría son de este tipo, pero existen también modelos más antiguos) en instalaciones europeas.

que las armas nucleares tácticas sean útiles en una guerra deben matar a sus víctimas lo más rápido posible. Una "incapacitación permanente inmediata" de acuerdo con pruebas recientes realizadas por el Gobierno de los Estados Unidos con monos rhesus, requiere 8000 rads. Puesto que los tanques modernos tienen un efecto protector de radiación de alrededor 0,5, éstos deben ser expuestos a 16.000 rads de radiación instantánea si se pretende alcanzar óptimamente los objetivos de la OTAN. No obstante, se ha revisado recientemente la doctrina militar de la OTAN para admitir que una "incapacitación temporal inmediata", que requiere sólo de 2500 a 3500 rads (o, si se tiene en cuenta la protección de los tanques, de 5000 a 7000 rads), puede ser suficiente a efectos militares para neutralizar la invasión.

Una persona expuesta a 8000 rads queda imposibilitada en cinco minutos y permanece inhábil para cualquier trabajo físico hasta su muerte, que deviene en uno o dos días. Una dosis de 3000 rads también incapacita en cinco minutos, pero la víctima puede recuperarse parcialmente a los 30 minutos, aunque está condenada a morir al cabo de cuatro o seis días. También puede sobrevivir y quedar incapacitado. (Esta incertidumbre tiene significativas implicaciones militares.) Una exposición a 650 rads deteriora funcionalmente un ser humano en dos horas, aunque puede responder a un tratamiento médico. Se parece a un doloroso y dilatado deterioro físico que terminará con la muerte en un par de semanas; ciertamente una horrible perspectiva, pero quizá suficiente plazo para que la víctima pueda seguir combatiendo durante algún tiempo.

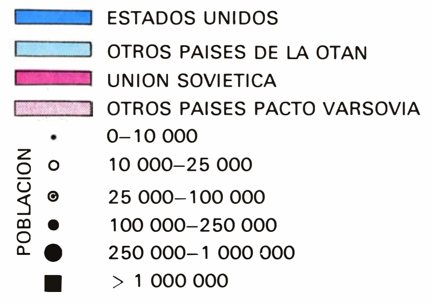
Estos resultados son consecuencia de los efectos ionizantes de los neutrones, que chocan con los protones en el interior de las células vivas. La ionización rompe los cromosomas, aumenta el núcleo celular y la viscosidad del fluido celular, refuerza la permeabilidad de la membrana de la célula y destruye células de todo tipo, particularmente las del sistema nervioso central. Además, la exposición a las radiaciones ionizantes demora o destruye el proceso de mitosis, efecto genético a largo plazo que inhibe la sustitución celular normal.

Las armas de radiación intensiva distribuyen una dosis dada de radiación sobre una superficie mayor que las bombas de fisión de potencia equivalente o incluso superior. Por ejemplo, una persona que se encuentra en el interior de un radio de 375 metros de una explosión de fisión de un kilotón (o de un radio de 630 metros de una explosión de fisión de 10 kilotones) estará expuesta a 8000 rads por lo menos. Pero si en vez de ello explosionara una bomba de radiación intensiva de un kilotón, el círculo de 8000 rads alcanzaría un radio de 850 metros. Así pues, una bomba de radiación intensiva de un kilotón puede matar potencialmente cerca del doble de conductores de tanques que una bomba de fisión de 10 kilotones, aunque el área devastada por la explosión alcanzaría sólo la quinta parte de su amplitud.

Desde la perspectiva de los oficiales de la OTAN, esta característica es obviamente el argumento principal a favor de la bomba de radiación intensiva. El punto clave es que la nueva arma puede reducir sustancialmente los perjuicios colaterales de una explosión nuclear, queriendo decir con esto que los efectos de la onda de choque, de la radiación térmica y de la lluvia radiactiva serán menos dominantes. Esto suena bien en un principio. No obstante, resulta erróneo pensar, como algunos de los defensores de la bomba parecen hacerlo, que con esta nueva generación de armas nucleares tácticas una guerra nuclear en territorio europeo podrá ahora ser manejada más segura y fácilmente de lo que en el pasado se creyó posible.

Con todo, una "guerra nuclear limitada" exige el consenso de las dos partes, y los rusos no parecen estar preparados ni dispuestos a llegar a un acuerdo de este tipo. Se cree que la mayoría de las 3500 bombas nucleares tácticas que tienen emplazadas contra objetivos europeos (en comparación con las 7000 de la OTAN), poseen una potencia superior a los 20 kilotones, y cerca de 600 misiles rusos tienen una potencia comprendida entre los 500 kilotones y tres megatones. Los proyectiles nucleares tácticos de las naciones del Pacto de Varsovia carecen de la precisión de los de la OTAN, lo cual hace difícil sino imposible llevar a cabo con éxito una táctica de impactos selectivos, necesaria para una estrategia eficaz de daños limitados. La doctrina militar rusa no parece reconocer ninguna fina distinción entre los tipos diferentes de

guerra nuclear táctica, como hacen a menudo los militares del estado mayor norteamericano. Verdaderamente, la mayoría de los manuales rusos al respecto dan por sentado la inevitabilidad de la escalada, y no hacen distinción entre una guerra nuclear táctica y una guerra nuclear estratégica completa. Cuando tales artículos tratan de una guerra nuclear en territorio europeo, no mencionan prácticamente la determinación selectiva y altamente precisa de objetivos, sino es quizá para ridiculizarlos. Una barrera de fuego masivo que perfore amplios huecos en las defensas de la OTAN, seguida por una invasión con tanques pesados (cuya estructura y material de superficie proporcione una cierta protección contra los efectos nucleares), parece ser el tipo de misión previsto para las armas nucleares tác-



INVASION HIPOTETICA de Europa Occidental por las fuerzas de los países del Pacto de Varsovia. Las áreas en color sombreadas irregularmente muestran el despliegue aproximado de las fuerzas de la OTAN y las del Pacto de Varsovia en el día de la movilización *(véase la clave para el código de color de las áreas de despliegue).* **Las fuerzas serían redesplegadas entre el día de la movilización y el día D (el primer día de guerra). Las flechas indican las principales rutas de invasión que se consideran como más probables. Los tres ejes verosímiles de ataque coinciden con las autopistas principales. Un intento de las fuerzas de la OTAN para detener las divisiones armadas rusas que dirigen cada uno de estos ataques mediante el lanzamiento de bombas nucleares de radiación intensiva requeriría una barrera de fuego de centenares de bombas muy adentrada en la frontera de Alemania Occidental. A causa de la urbanización creciente de la región, un contraataque de este tipo podría matar a varios cientos de miles, e incluso varios millones, de civiles y combatientes de la OTAN (sin incluir las muertes que produciría la probable respuesta nuclear rusa y de las fuerzas del Pacto de Varsovia). La distribución de la población se indica por símbolos clave.**

ticas desde el punto de vista de la Unión Soviética.

Si la OTAN utilizara bombas de radiación intensiva contra los tanques del Pacto de Varsovia, los rusos responderían casi con toda certeza con sus propias armas nucleares. Tal como afirma un estudio de la inteligencia militar norteamericana sobre las operaciones militares rusas: "En caso de que el primer escalón [de tanques en una ofensiva] fracasara, seguiría una serie de contraataques coordinados con todas las unidades de combate hasta incluir... impactos nucleares". Los rusos no prestarían probablemente mucha atención a los daños colaterales que sufriera la población civil de Alemania Occidental; y aún cuando la prestaran, la alta potencia y la escasa precisión de sus armas les impediría actuar eficazmente para evitar tales efectos.

Es más, antes de la casi segura respuesta nuclear rusa, el daño causado por las bombas de radiación intensiva de la OTAN sería notable, a pesar de las supuestas limitaciones en la onda de choque, en la radiación térmica y en la lluvia radiactiva de cada bomba individual. El informe del Departamento de Defensa de los Estados Unidos establecía, para el año fiscal 1977, que si se utilizaran armas nucleares en Europa, tal acción "...debería inducir a la Unión Soviética a terminar rápidamente el conflicto... Debería realizarse con decisión y contundencia para conseguir que los soviéticos reconsideraran sus actividades". Para alcanzar tal contundencia, la OTAN debería hacer algo más que detener un pequeño número de tanques; se necesitará ciertamente mucho más destrucción para aportar un cuadro dramático a los dirigentes rusos.

¿Qué contundencia debería alcanzarse? Al empezar una ofensiva los tanques rusos se moverían en dos escalones (o en tres según las circunstancias). Los tanques del primer escalón irían separados unos 75 metros en las acciones no nucleares y 100 metros en las nucleares. El segundo escalón avanzaría a unos tres kilómetros de distancia del primero. El Pacto de Varsovia tiene unos 20.000 tanques desplegados en la región central de Europa, en donde tal vez podría ocurrir la primera batalla de la guerra OTAN/Pacto de Varsovia. Las afirmaciones de los oficiales del Ejército norteamericano de que la bomba de radiación intensiva produce pequeños daños colaterales tiene validez si se corresponde con una alta selectividad, incluso a nivel individual, de los impactos de las armas utilizadas. Con lo cual, si la OTAN quiere detener una fracción notable del primer escalón de tanques, es decir, si las armas de radiación intensiva deben ser útiles militarmente, la acción exigirá un bombardeo de muchos cientos e incluso miles de armas nucleares. Con suma probabilidad no sólo incluiría bombas de radiación intensiva de baja potencia, sino también bombas de fisión de baja y media potencia. En tales circunstancias caería mucha radiactividad en el suelo, especialmente si algunas de las bombas detonaran por accidente sobre o cerca de la superficie. En cualquier caso, el número de fatalidades y accidentes debidos a "espíritus caminantes" irradiados sería muy elevado, aunque la guerra nuclear se mantuviera en unos límites restringidos.

Las bombas de radiación intensiva pueden reducir el daño colateral causado por la onda de choque y la radiación térmica, pero los efectos causados por la radiación instantánea aumentarán. La exposición a la radiación, incluso en dosis relativamente pequeñas, puede tener graves consecuencias para los seres humanos, y las bombas de radiación intensiva incrementarían el área en la cual la gente recibiría dosis peligrosas. Por ejemplo, el 10 por ciento de la gente expuesta a 150 rads moriría de enfermedades derivadas de las radiaciones; (los sobrevivientes de Hiroshima y Nagasaki expuestos a 150 rads mostraron una extraordinariamente alta incidencia de cáncer de pecho). La exposición a sólo 30 rads dobla la tasa de mutaciones en los descendientes, y es de esperar la aparición de genes deficientes hasta dentro de 10 generaciones. Los habitantes de las islas Marshall que estuvieron expuestos a sólo 14 rads, a consecuencia de las pruebas nucleares de los Estados Unidos en 1954, sufrieron recientemente nódulos de tiroides, cáncer y leucemia.

Una bomba de radiación intensiva de un kilotón libera 150 rads a una distancia de 1,7 kilómetros, 30 rads a 2,1 kilómetros y 14 rads a 2,3 kilómetros. Hay que comparar estas distancias con 900, 1170 y 1300 metros, respectivamente, para una bomba de fisión de un kilotón, y con 1285, 1570 y 1700 metros para una bomba de fisión de 10 kilotones.

Los efectos colaterales de una bomba de radiación intensiva causados por la radiación instantánea podrían ser incluso mayores. Se cree que para la radiación de los rayos gamma existe un umbral por debajo del cual no se producen daños biológicos, umbral que no se acepta para la radiación de neutrones. Además, si atendemos a los daños genéticos, leucemia y cataratas, los efectos biológicos de los neutrones son cerca de seis veces más elevados que los de los rayos gamma. De este modo, basta un rad o dos de radiación de neutrones para causar leucemia y cáncer. Exposiciones a sólo cinco rads pueden doblar la tasa de mutaciones en los descendientes de las personas expuestas. Si un solo neutrón colisiona con una hebra de ADN de una célula del espermatocito o de óvulo, es alta la probabilidad de daño genético irreparable a largo plazo. En otras palabras, la idea de que las armas de radiación intensiva son claramente benignas a la gente de "nuestro lado" resulta muy cuestionable. Tanto los combatientes como los no combatientes de la OTAN pueden sufrir graves daños. El riesgo para los no combatientes viene incrementado por el hecho de que las tierras orientales de Alemania Occidental están densamente urbanizadas.

El propio interés militar resulta cuestionable a un nivel más elemental. Excepto para los hombres que conducen los tanques, que se hallarían en las inmediaciones del lugar de explosión, el resto de tropas enemigas expuestas podría permanecer a salvo durante horas, días o incluso semanas; los soldados podrían combatir de un modo más agresivo que antes, puesto que tendrían la certeza de su próxima muerte por efecto de la radiación. Desde luego, la OTAN podría resolver este problema arrojando un mayor número de armas de radiación intensiva. Puesto que las ventajas alegadas a propósito de la bomba de radiación intensiva descansan, principalmente, en su capacidad de realizar impactos limitados, selectivos y precisos, este tipo de bombardeo masivo trunca de raíz la racionalidad del arma. Además, los neutrones que penetren en el blindaje no dejarían el tanque con la suficiente radiactividad como para evitar que otra tripulación reemplace a la anterior, víctima de la radiación. Los tanques podrían continuar avanzando.

El uso eficaz de esas armas presupone también una concentración masiva de tanques. Con todo, puede afirmarse con seguridad que las naciones de la OTAN no darían orden de fuego de armamento nuclear hasta que las naciones del Pacto de Varsovia no hubieran castigado y atravesado las defensas no nucleares de la OTAN. Aunque los rusos hubieran concentrado sus tanques en las fases iniciales de la ofensiva, podrían dispersar ciertamente sus fuerzas blindadas después de romper las defensas fronterizas de la OTAN. (Sus manuales de operaciones tácticas sugieren que esto es exactamente lo que harían.) Bajo estas condiciones deberían utilizar-

se muchos miles de bombas de radiación intensiva a fin de incapacitar inmediatamente a los ocupantes de un número significativo de tanques del Pacto de Varsovia, anulando de nuevo las alegadas virtudes de este tipo de bomba.

A pesar de su aparentemente mínima utilidad militar, el Departamento de Defensa de los Estados Unidos justifica las bombas de radiación intensiva en los siguientes términos: "si los arsenales de la OTAN tuvieran la bomba de neutrones, los países enemigos tendrían en cuenta la capacidad de la OTAN para defenderse con menos pérdidas propias: esto *podría constituir* un medio disuasor de un posible ataque". Aunque el Departamento de Defensa no establece explícitamente que esta arma podría reforzar la disuasión, se entiende que los rusos tendrán en cuenta que la OTAN utilizará con mayor facilidad las armas de radiación intensiva que las antiguas bombas nucleares.

Este punto merece tres apostillas. (1) Incluso sin la amenaza de las armas de radiación intensiva, el riesgo que asumirían los rusos al iniciar un ataque sería sumamente elevado, puesto que los Estados Unidos se han negado firmemente a adoptar una política de reacción ante otro primer usuario del armamento nuclear. (2) El uso de las armas de radiación intensiva causaría grandes perjuicios a la OTAN, y nada digamos de los procedentes de la obvia respuesta nuclear soviética. (3) Aunque este tema resulta demasiado complejo para tratarlo aquí, no hay razón para pensar que la OTAN sea incapaz de defender Europa Occidental sin acudir a las armas nucleares. Los niveles de potencia de fuego convencional entre las naciones de la OTAN y las del Pacto de Varsovia son prácticamente equivalentes, y es una regla bien conocida que el atacante necesita una superioridad sustancial. La repetida superioridad de las naciones del Pacto de Varsovia en número de tanques viene contrarrestada por la ventaja que detenta la OTAN en armamento anticarro, especialmente después de las recientes innovaciones en municiones guiadas con precisión ("precision-guided munitions") y en vehículos por guía remota ("remote-guidance vehicles"). Estos dos nuevos tipos de armas tienen un alcance mayor que los cañones de los tanques rusos, y ambos pueden, según las palabras recogidas de un manual de campo del ejército norteamericano, "alcanzar lo que ven y matar lo que alcanzan".

El adiestramiento militar en la Unión Soviética y en los demás países de Europa Oriental es notoriamente pobre y extremadamente rígido. La confianza política de las divisiones checoslovacas y polacas, por lo menos en una guerra ofensiva, es dudosa. La táctica y la estrategia de los ejércitos del Pacto de Varsovia descansan pesadamente en el tanque, el cual se está convirtiendo en un sistema de armamento crecientemente vulnerable y obsoleto. Por otro lado, los numerosos escenarios de ataques por sorpresa que abundan en estos días no tienen en cuenta los bajos niveles de agilidad de las tropas del Pacto de Varsovia, los centenares de vías por las que los servicios de inteligencia pueden observar y detectar signos de movilización, la deficiente red logística rusa y otras muchas debilidades de la máquina de guerra soviética.

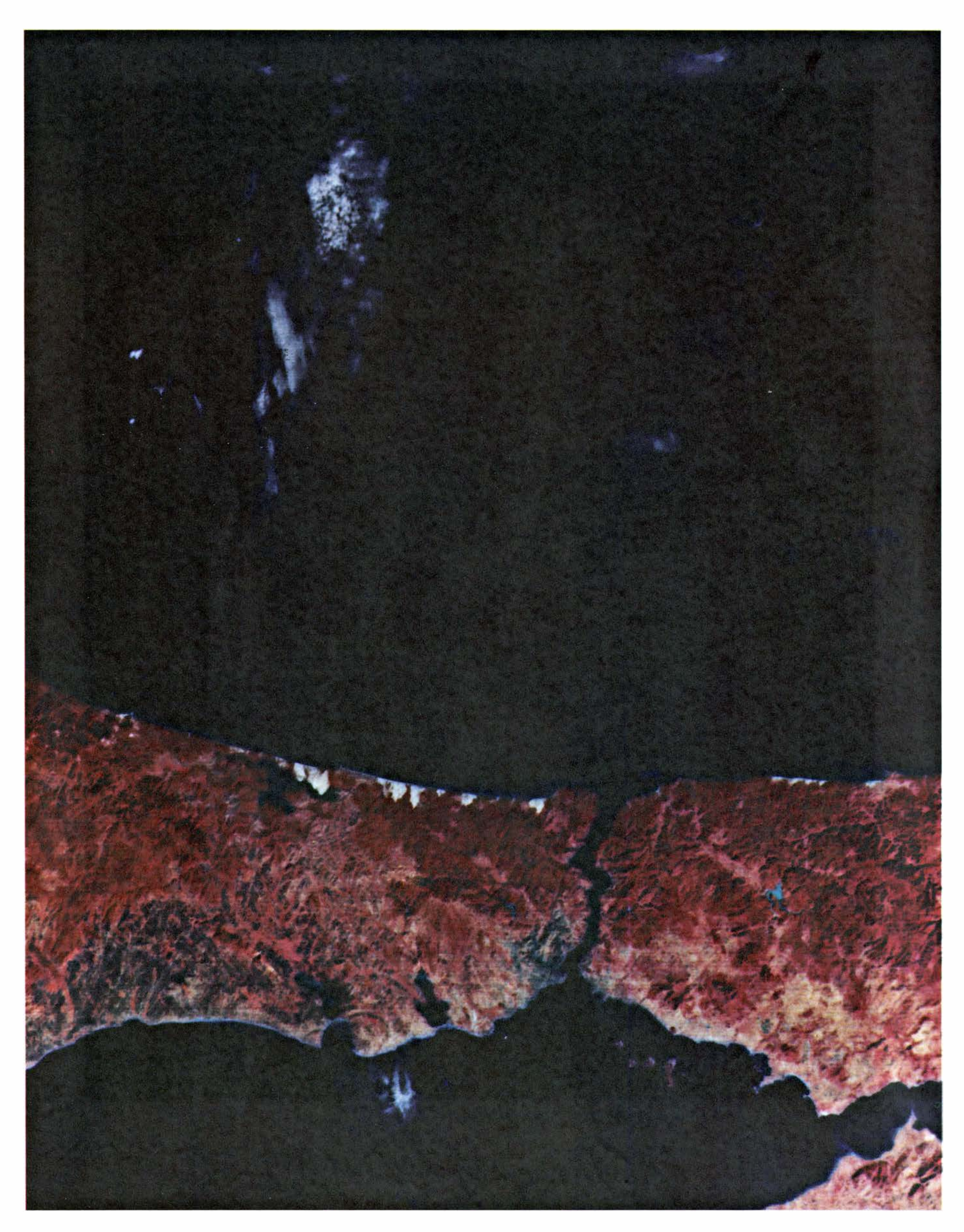
Esto no significa que no haya nada que mejorar en la OTAN. Varios emplazamientos incorrectos de fuerzas pueden corregirse; las líneas de comunicación pueden desplazarse hacia atrás, lejos del eje previsto del campo de batalla; podrían desplegarse más armas convencionales anticarro; los aeropuertos podrían estar más dispersos; podrían emplazarse más aviones en los lugares "endurecidos". Hay indicios de que la actual Administración de los Estados Unidos va a afrontar estos problemas. Y como quiera que algunos de estos puntos requieren gastos muy cuantiosos, parece descabellado asignar grandes sumas de dinero a estos fines al mismo tiempo que a la bomba de radiación intensiva.

Los costes de producción de las bombas de radiación intensiva serían enorme. El obús de ocho pulgadas costaría unos 900.000 dólares por unidad (incluyendo el coste del proyectil, la cubierta y todo lo demás). Para la versión en radiación intensiva del misil Lance se calcula un coste sólo ligeramente inferior. En lugar de adquirir dos andanadas de bombas de radiación intensiva para obuses de ocho pulgadas, los Estados Unidos podrían obtener, digamos, tres tanques de combate abierto de M-60, unas 50 armas anticarros no nucleares avanzadas, o más de 5500 andanadas de bombas convencionales de artillería. En otras palabras, si los Estados Unidos deciden invertir en armas de radiación intensiva, la OTAN adquirirá un arma extraordinariamente costosa que quizá nunca llegará a emplear, a expensas de armas más baratas en comparación que podrían mejorar notablemente su cuadro defensivo. En el supuesto de que los rusos dispersen sus tanques ampliamente y adopten ciertas medidas contra la radiación de neutrones, el armamento anticarro convencional resultaría probablemente más barato y de mayor eficacia militar.

Continúa siendo válido que la bomba de radiación intensiva puede hacer tanto daño a una fuerza atacante como otras bombas de mayor potencia sin provocar tanta destrucción colateral. Sin embargo, en contra de esta clara ventaja, deben tenerse en cuenta las graves consecuencias que puedan resultar de la introducción de armas nucleares en una guerra clásica. También podría argumentarse a favor de las bombas de radiación intensiva que, por tratarse de armas de una nueva generación, incluyen una serie de dispositivos modernos que mejoran los rendimientos de las anteriores.

Tendrían un alcance mayor (cerca de 130 kilómetros para la versión nueva del misil Lance), sistemas de comunicación de control y guía mejorados y mecanismos de mayor seguridad. Estos nuevos dispositivos tendrían un efecto estabilizador en el sentido de que harían menos vulnerables a las armas nucleares tácticas ante un ataque convencional del Pacto de Varsovia, y menos susceptibles de un disparo accidental. Tales dispositivos pueden incorporarse fácilmente a la generación actual de armas nucleares tácticas; la radiación intensiva no es necesaria para tales propósitos.

Particularmente peligrosa resulta la bomba de radiación intensiva en cuanto induce a creer erróneamente que su desarrollo hará posible una guerra nuclear limitada y controlada; en este sentido, su propio despliegue puede hacer bajar el umbral que separa la guerra clásica de la nuclear. Las armas de radiación intensiva no son más (y quizá sean menos) "humanas" que las armas químicas, cuyo primer uso fue proscrito hace tiempo por tratados internacionales. Además de poseer una escasa superioridad militar sobre cualquier otro tipo de bomba nuclear de baja potencia. Finalmente, en el supuesto de que la Unión Soviética crea que los Estados Unidos utilizarán armas de radiación intensiva en una guerra europea, su despliegue constituye una invitación a un ataque nuclear preventivo ruso en una situación extremadamente tensa, y quizá como primer movimiento bélico. En cualquier caso, no hay razón para creer que la bomba de radiación intensiva pueda disminuir de algún modo la probabilidad de escalada de una guerra nuclear en territorio europeo hacia una guerra nuclear total, ni que su introducción pudiera moderar de alguna manera la probable respuesta soviética.



EL ESTRECHO DEL BOSFORO es la única comunicación actual entre el mar Negro y el Mediterráneo. El mar Negro se encuentra en la parte superior; la masa de agua de la parte inferior es el mar de Mármara, que comunica con el Mediterráneo a través de los Dardanelos, situados a unos 500 kilómetros hacia el sudoeste. El Bósforo tiene una longitud de unos 30 kilómetros y durante casi todo este recorrido su anchura es de uno o dos kilómetros. Su profundidad es también bastante pequeña. En tiempos pasados, con el nivel de los mares más bajo, el Bósforo era un río que drenaba un lago de agua dulce situado en la cuenca del mar Negro. La región obscura del extremo meridional del Bósforo es Estambul. La fotografía se tomó desde el satélite Landsat 2. Se trata de una imagen en falso color donde la vegetación aparece en rojo intenso.

Cuando el mar Negro se desecaba

No hace seis millones de años, los ríos que desembocaban en el mar Negro fueron desviados hacia la cuenca desecada del Mediterráneo. Tras un período de desecación, el mar Negro fue, hasta época reciente, un lago de agua dulce

Kenneth J. Hsü

Ni siquiera los rasgos más prominentes y de mayores dimensiones de la superficie terrestre pueden considerarse permanentes. Puesto que las placas rígidas que constituyen la superficie están dotadas de movimiento, los continentes y océanos se ven constantemente remodelados, sufriendo cambios en sus disposiciones relativas. La formación de montañas, asociada a los movimientos de las placas, altera las redes de drenaje y el clima, produciéndose a su vez, necesariamente, una respuesta por parte de la vida animal y vegetal de las regiones afectadas. La geología histórica da cuenta de varios ejemplos de este tipo; otros pueden observarse hoy en fase de desarrollo, con la ventaja de que la tectónica de placas, teoría formulada para explicar esos cambios, se halla en la actualidad bien consolidada. A pesar de ello no deja de sorprender el que un determinado rasgo familiar de nuestro globo, que ha presentado la misma forma desde la aparición de los primeros mapas, haya estado sujeto a súbitas transformaciones.

En 1970 se descubrió un extraordinario episodio de este tipo. Los testigos de los sedimentos extraídos del fondo del mar Mediterráneo revelaron que toda esta extensión de agua se secó hace unos seis millones de años. Durante casi un millón de años, el fondo marino se convirtió en una cuenca desértica situada dos kilómetros por debajo de los altiplanos continentales circundantes. Luego, la cuenca quedó invadida de nuevo por las aguas procedentes del Atlántico; el relleno tardó varios siglos en producirse y, durante este tiempo, el estrecho de Gibraltar debió funcionar como la más espectacular de las cataratas.

Se ha podido saber que el mar Negro tiene una historia cuyo desarrollo ha sido muy parecido al experimentado por el mar Mediterráneo. Inicialmente, el mar Negro era un brazo de un gran océano; más tarde pasó a formar parte de un gran mar interior que cubría una vastísima extensión de Europa oriental. Debido a una breve crisis se desecó casi por completo y, después de haberse llenado de nuevo, se convirtió en un profundo lago de agua dulce. El mar de aguas salobres y eminentemente estancadas que en la actualidad ocupa la cuenca comenzó su evolución en las postrimerías de las glaciaciones, y su edad no es mucho mayor que la de las civilizaciones que se desarrollaron en sus cercanías por la parte de Levante.

Hoy en día, el mar Negro tiene una longitud aproximada de 1200 kilómetros y una profundidad máxima de unos 2200 metros. Es la cuenca que drena una parte importante del continente europeo y recibe varios ríos principales, entre los que están el Danubio, el Dniester, el Dnieper y el Don. Se comunica con el Mediterráneo a través del estrecho pasadizo constituido por el Bósforo, el mar de Mármara y los Dardanelos.

Tanto el mar Mediterráneo como el mar Negro son restos de un antiguo océano ecuatorial que separaba Africa de Europa y comunicaba el Atlántico con el océano Indico. En 1900, un geólogo vienés, Eduard Suess, llamó a este mar Tetis, derivado de Titán, de la mitología griega, hija de Gea y esposa de Oceano. Hace unos 20 millones de años, el movimiento hacia el norte de la placa africana empezó a cerrar el mar Tetis. La colisión con Asia cerró la comunicación oriental con el océano Indico. La colisión con Europa produjo una larga cordillera de montañas que comprende los Alpes, los Dinárides y Helénides de Yugoslavia, Albania y Grecia y el Taurus de Turquía meridional. Como resultado de la formación de estas montañas se dividió el Tetis en dos mares interiores. Uno era el antiguo Mediterráneo, que ocupaba la misma cuenca que en la actualidad. El otro mar, situado al norte y al este, ha recibido el nombre de Paratetis.

El Paratetis se extendía desde la cuenca húngara, al oeste, a través de Europa sudoriental y de lo que es hoy el mar Negro y el mar Caspio hasta más allá del mar de Aral, en la Rusia asiática. La comunicación entre el Paratetis y el Mediterráneo se cortó hace unos 15 millones de años. Tras lo cual, el Paratetis se convirtió en un mar de aguas salobres muy parecido al Báltico actual. El geólogo francés Maurice I. M. Gignoux sugirió, en 1920, que fuera llamado Lac Mer, literalmente Lago Mar, para poner de manifiesto que era intermedio entre un lago y un océano abierto.

Las únicas masas de agua importantes del Lac Mer que han sobrevivido hasta nuestros días son el mar Negro, el mar Caspio y el mar de Aral. Se ha visto ya claramente que la desintegración de esta especie de gran canal interior estuvo estrechamente asociada con la súbita desecación del Mediterráneo. El registro detallado de algunos, al menos, de los acontecimientos de esta historia se ha ido revelando en la lectura de los sedimentos depositados sobre el fondo del mar Negro.

La catastrófica desecación del Mediterráneo se descubrió al sondear centenares de metros de rocas sedimentarias y barros del fondo marino y extraer los correspondientes sedimentos en forma de testigos de sondeo. Los sondeos fueron realizados por el buque de investigación *Glomar Challenger*, especialmente construido para realizar trabajos de este tipo en zonas de mar profundo. Sobre la cubierta del buque se levanta una gran torre de perforación, y existe la posibili-

dad de utilizar y almacenar más de siete kilómetros de tubos de perforación. Igualmente importantes son los aparatos destinados a la navegación y un sistema de propulsión especial que mantiene la posición del barco exactamente encima del sondeo, en medio incluso de un temporal. Botado en 1968, el *Glomar Challenger* está al servicio de la Scripps Institution of Oceanography bajo el asesoramiento de un grupo internacional de investigadores que constituyen la comisión del Joint Oceanographic Institutions for Deep Earth Sampling (JOIDES).

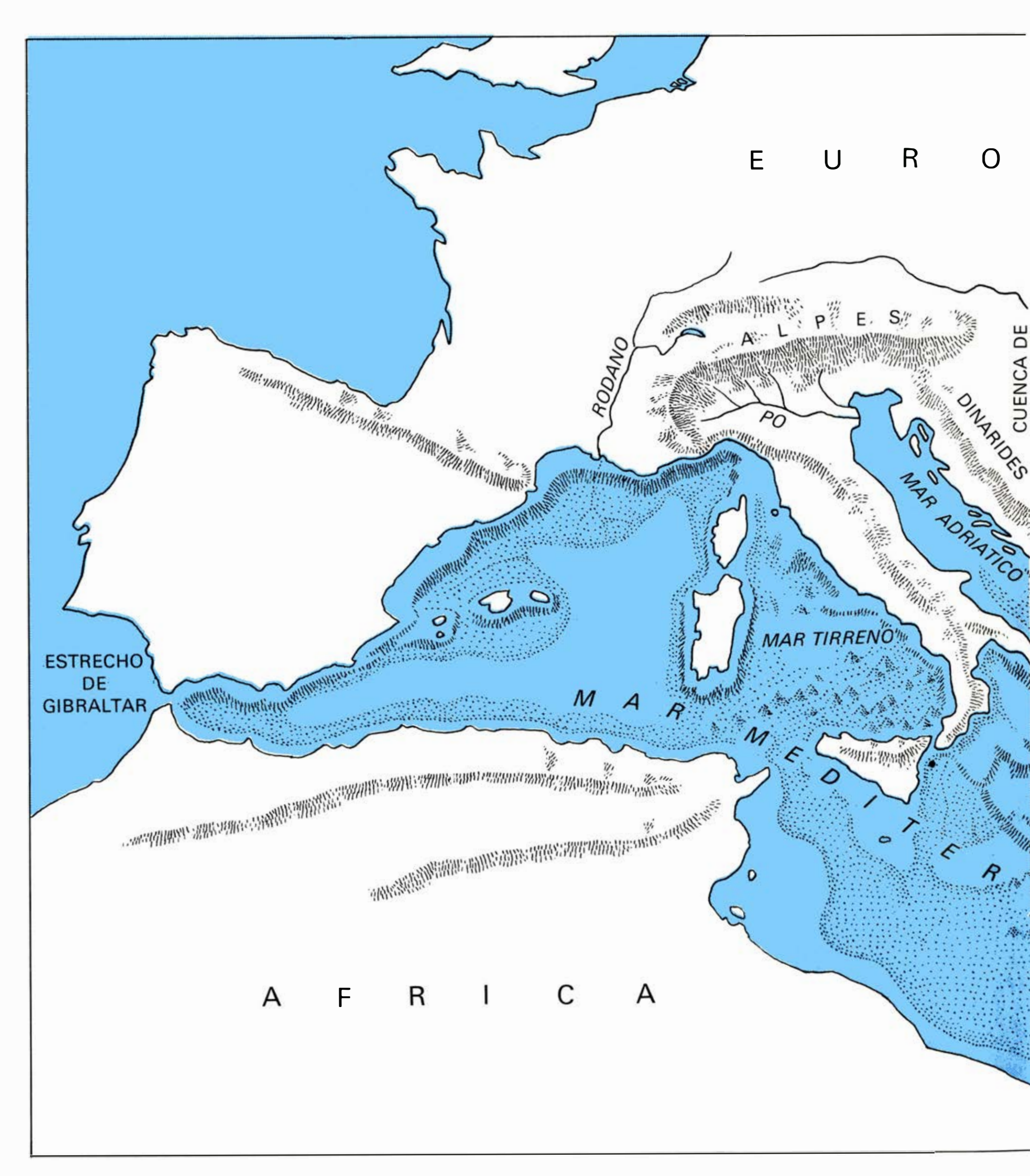
Los primeros estudios sísmicos realizados en el Mediterráneo habían revelado un horizonte de material situado a unos 100 metros por debajo del fondo marino que actuaba como un fuerte reflector de ondas sonoras. En 1970, el viaje del *Glomar Challenger,* en el cual yo participé, reveló la naturaleza de este horizonte reflector. Está formado por rocas carbonatadas y sulfatadas, que se depositan típicamente en aguas poco profundas, y por sales que precipitan gracias a la evaporación completa de las salmueras. Una posible explicación de la presencia de este horizonte era obvia e inmediata, aunque resultara también extrema: el horizonte pudo haberse formado a condición de que el Mediterráneo se hubiera quedado seco por evaporación. Después del viaje, mis colegas y yo mismo estábamos capacitados para reconstruir cómo pudo haber ocurrido tal cosa.

El Mediterráneo tiene un gran déficit hidrológico: de su superficie se evapora más agua que la que aportan las cuencas hidrográficas circundantes. Su nivel se mantiene solamente merced al gran volumen de agua aportada a través del estrecho de Gibraltar. Hace unos seis millones de años, el movimiento continuado hacia el norte de la placa africana produjo la formación de cordilleras en la península Ibérica y en el extremo noroccidental de Africa, creando un istmo que aislaba al Mediterráneo. A lo largo de los 1000 años siguientes, más o menos, el enorme volumen de este mar, unos cuatro millones de kilómetros cúbicos, se evaporó depositándose sales y otros residuos en la cuenca desecada. Hace poco más de cinco millones de años, el Atlántico penetró a través del estrecho de Gibraltar y la cuenca se convirtió de nuevo en un mar. El relleno de la cuenca es un dato que marca la división entre dos épocas del tiempo geológico: el Mioceno (menos reciente) y el Plioceno (reciente). Después de estas épocas vino el Pleistoceno (más reciente), hace aproximadamente 1,8 millones de años. La época actual es el Holoceno, que empezó hace unos 10.000 años.

Esta interpretación de la historia del Mediterráneo no se aceptó sin más ni más, y hubo un argumento, que aportó pruebas que parecían contradecir tal interpretación, particularmente sutil. En una muestra datada en el Mioceno superior y procedente de un sondeo del Mediterráneo oriental, encontramos unos fósiles consistentes en pequeños crustáceos pertenecientes al género *Cyprideis* de la subclase Ostracoda. Las especies de *Cyprideis* eran propias de agua dulce o de aguas salobres y no hubieran podido sobrevivir en salmueras y mucho menos en una llanura desértica. Evidentemente el Mediterráneo no estaba siempre seco a pesar de estar aislado; de vez en cuando las llanuras salíferas hundidas debían haber estado cubiertas por grandes lagos. ¿De dónde procedía el agua dulce necesaria para llenar esos lagos?

Determinadas especies que viven hoy en el Mediterráneo o en aguas circundantes habían representado igualmente un problema. En muchas cuevas submarinas a lo largo de la costa adriática de Yugoslavia, por ejemplo, existen especies relictas que parecen representar "fósiles vivientes", restos de una fauna antigua que sólo ha sobrevivido en esos hábitat tan aislados. Se conocen otras poblaciones relictas en regiones costeras del sur de Francia, España y norte de Africa; en 1891, se descubrió un grupo especialmente notable de especies de este tipo en el lago Ohrid, que es un lago de montaña situado en la frontera entre Yugoslavia y Albania.

Por definición, las poblaciones relictas no están relacionadas con las existentes en las aguas que las rodean, pero las



LOS MARES ACTUALES de Europa sudoriental y de Asia sudoccidental son restos de un antiguo mar ecuatorial, llamado Tetis, que comunicaba el Atlántico con el océano Indico. Las principales masas de agua supervivientes son el Mediterráneo, que se sigue comunicando con el Atlán-

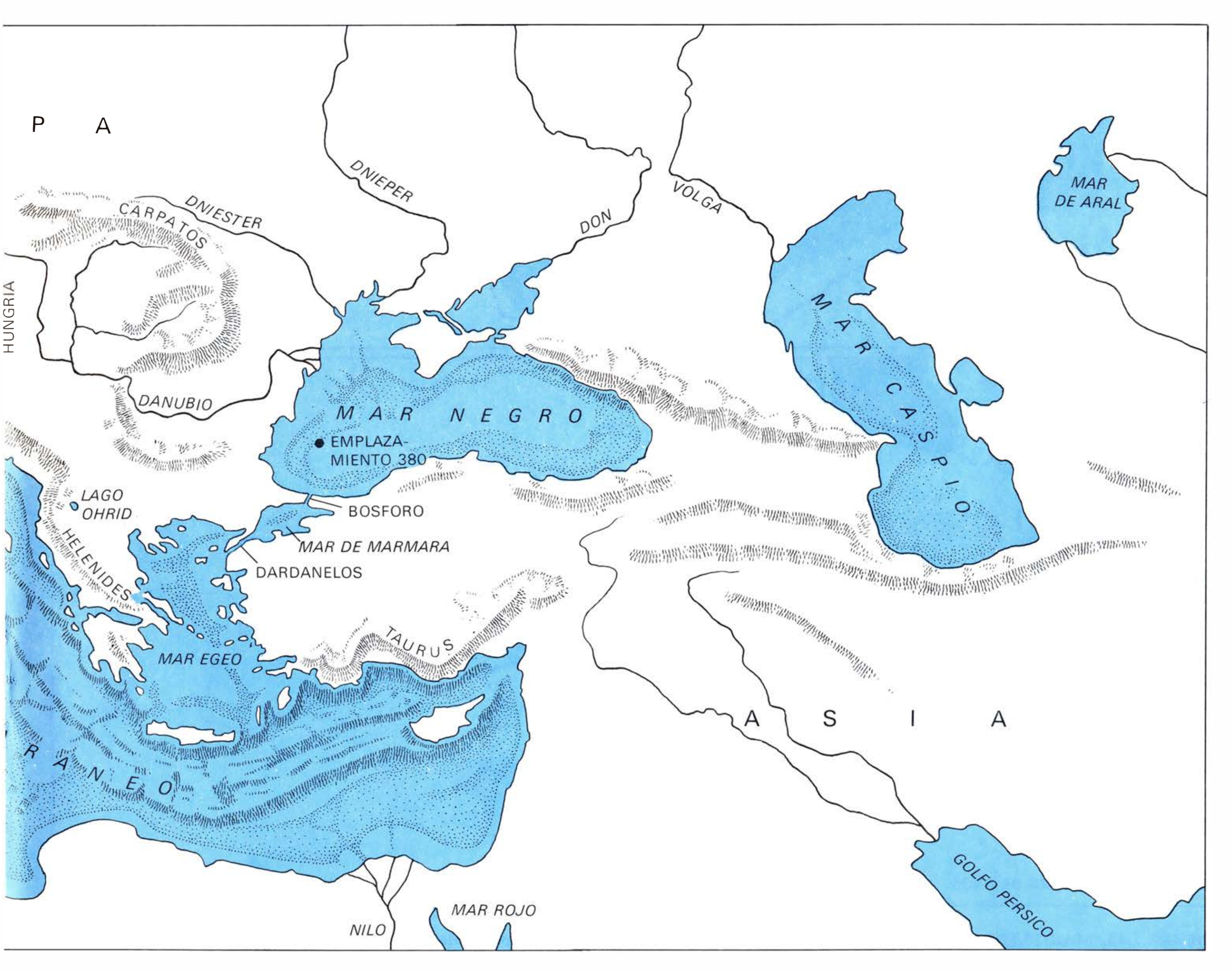
especies de cada una de estas zonas aisladas están relacionadas con especies de otras zonas parecidas. Más destacable todavía, muchas de ellas tienen especies afines en el distante, y completamente aislado, mar Caspio. Todas poseen una ascendencia común, que puede seguirse en el registro fósil del Paratetis mejor que en el del mar Mediterráneo.

Así pues, había bastantes integrantes sueltos que debían tenerse en cuenta, pero en realidad se trataba de un caso donde la multiplicación de misterios dio, como producto final, una hipótesis simple. La desintegración del Lac Mer, la presencia de lagos de agua dulce en la cuenca mediterránea y la dispersión de las especies del Paratetis eran hechos que podían explicarse si se suponía que el Paratetis había desaguado en el Mediterráneo en algún momento del Mioceno superior. La forma más inmediata de comprobar la hipótesis era examinar el registro sedimentario de una de las cuencas residuales del Paratetis.

En 1973 fui responsable de un grupo de expertos requerido por el JOIDES para dictaminar sobre una serie de propuestas de sondeo en el mar Negro. Se decidió que los sondeos podrían clarificar adecuadamente las relaciones existentes entre el Mediterráneo y el Paratetis, planificándose otro viaje del *Glomar Challenger*.

Otra razón que justificaba los sondeos en el mar Negro era estudiar el origen del petróleo. Por debajo los 200 metros de profundidad, el agua del mar Negro se halla estancada; no contiene oxígeno y está enriquecida en el tóxico gas sulfhídrico, de modo que las zonas profundas del fondo no pueden mantener otras formas de vida que no sean bacterias anaerobias. Bajo estas condiciones, la materia orgánica tiende a conservarse. Se cree que el petróleo se forma cuando los sedimentos orgánicos, así preservados, son enterrados y convertidos en hidrocarburos a temperatura y presión elevadas.

Otro de los fines del proyecto consistía en ayudar a clarificar el registro existente de la historia climática durante las glaciaciones. Generalmente se admiten cuatro períodos en la glaciación europea, que se denominan Günz, Mindel, Riss y Würm, nombres que provienen de los lugares de Alemania meridional donde se describieron por primera vez las formaciones características. Se han identificado igualmente cuatro períodos de glaciación en América del Norte que se denominan Kansan, Nebraskan, Illinoian y Wisconsin. La cuestión de si las dos cronologías se corresponden no ha sido

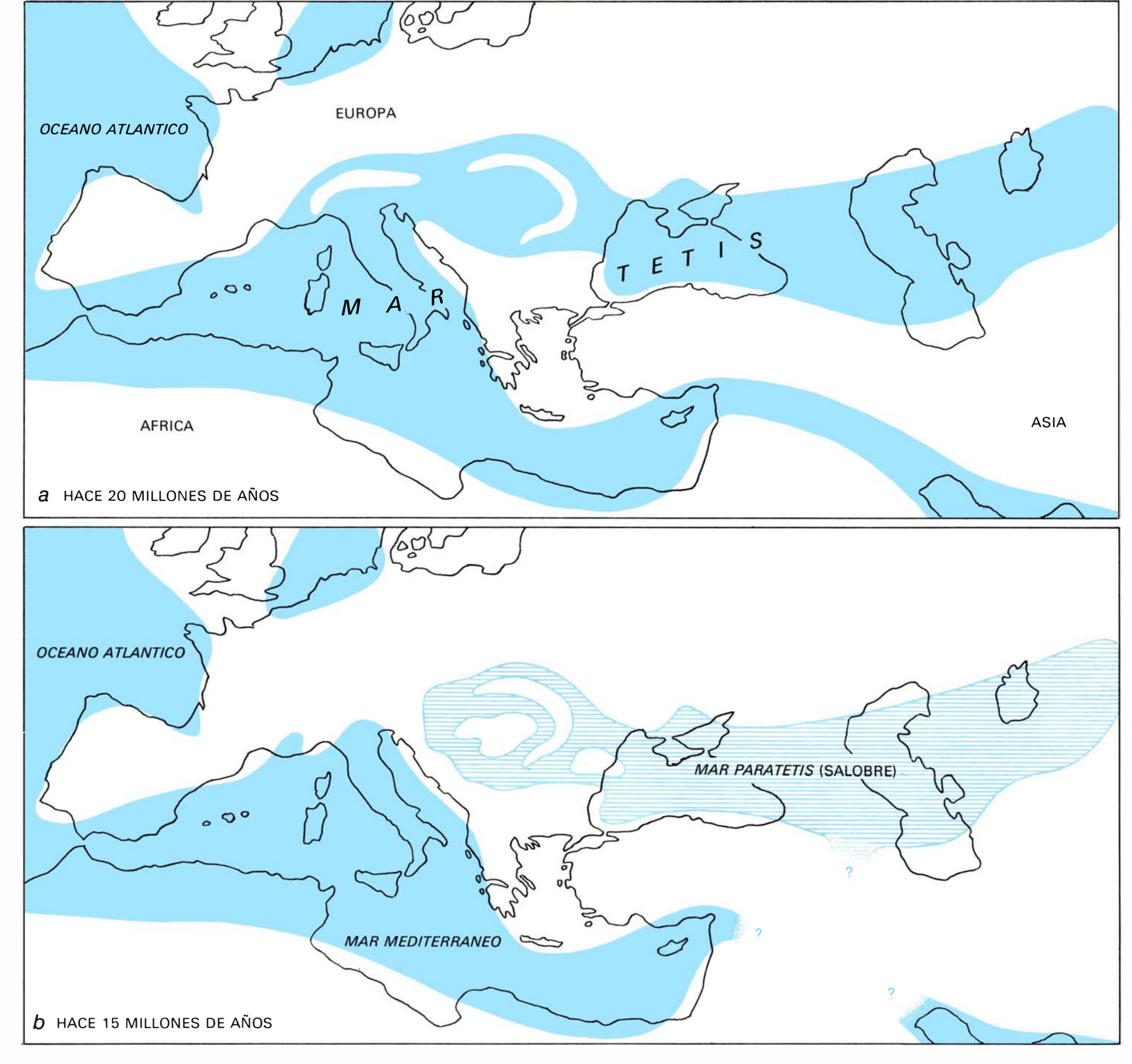


tico a través del estrecho de Gibraltar; el mar Negro, con una débil comunicación con el Mediterráneo a través del Bósforo, y los mares Caspio y de Aral actualmente aislados. La historia de los mares ha sido investigada mediante testigos de sondeo obtenidos hace muy poco tiempo en los sedimentos del fondo del mar Negro. El testigo de mayor longitud se extrajo del Emplazamiento 380, cerca del talud continental oeste.

nunca resuelta debido, sobre todo, a que no existe un registro continuo de toda la época glacial en ninguna serie sedimentaria que aflore en tierra firme. Los sedimentos marinos sí proporcionan un registro completo, pero su interpretación es ambigua. Algunos geólogos distinguen en ellos cuatro grandes períodos, comparables a los existentes en tierra firme, pero otros investigadores han detectado la existencia de muchos avances glaciares con una duración media de 100.000 años, o menos. Ni siquiera se ha podido datar con precisión el inicio de las glaciaciones; puede coincidir con el comienzo del Pleistoceno, hace unos 1,8 millones de años, pero podría haber ocurrido también mucho antes o mucho después.

El *Glomar Challenger* entró en el Bósforo el 19 de mayo de 1975 con una tripulación de 50 personas y un equipo científico de otras 20, dirigido por David A. Ross, de la Woods Hole Oceanographic Institution, y por Yuri Neprochnov, del Instituto de Oceanografía de Moscú. En un viaje de sólo tres semanas se realizaron sondeos en tres lugares. El sondeo más profundo se realizó en la base del talud continental occidental, frente a la costa de Bulgaria, en un lugar denominado Emplazamiento 380. Allí, con una profundidad de agua de 2000 metros, se obtuvo un testigo de 1073 metros de longitud.

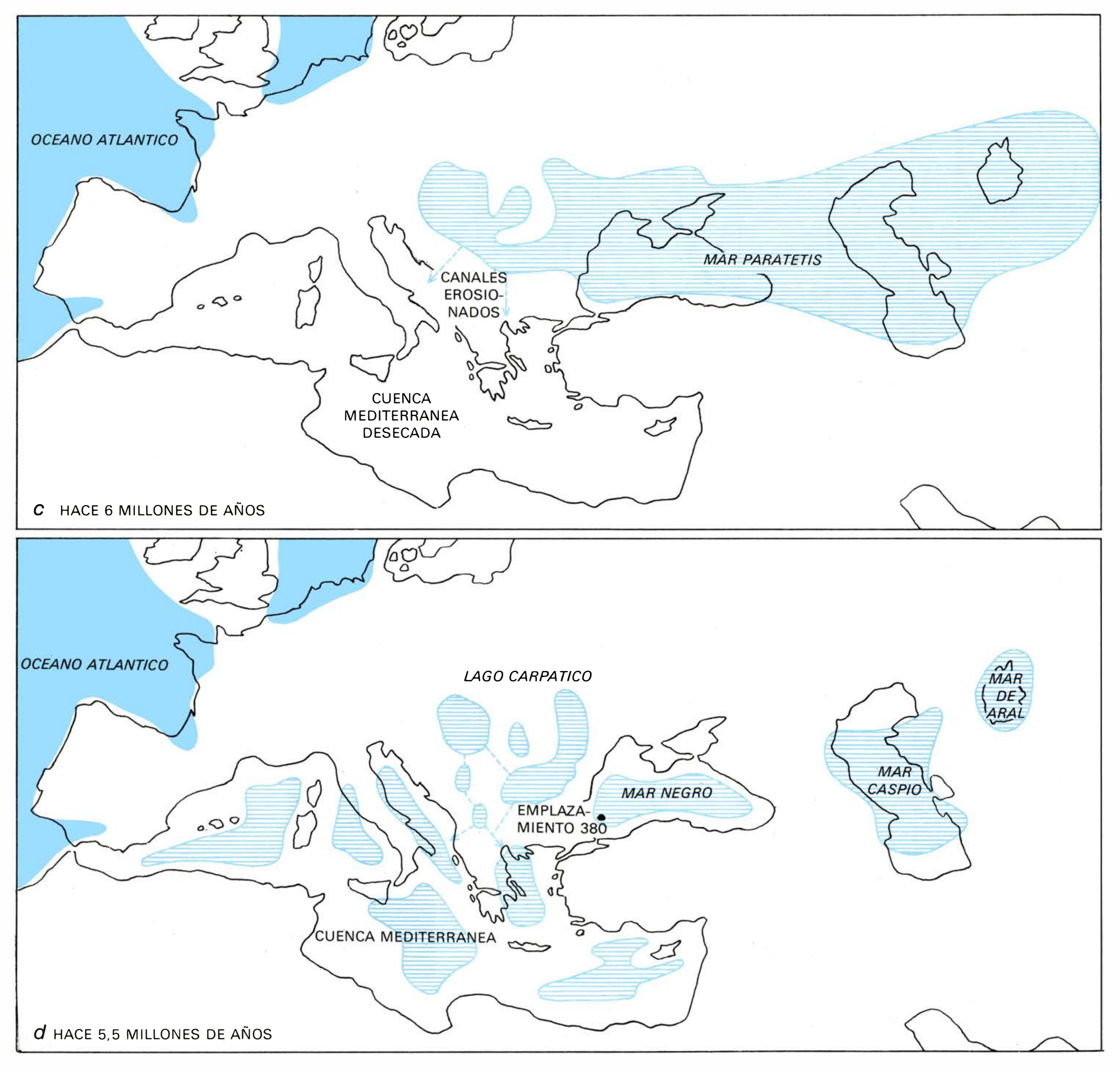
En las primeras inspecciones del testigo se vio claramente que las aguas marinas no habían ocupado siempre el Bósforo. En realidad, antes incluso de



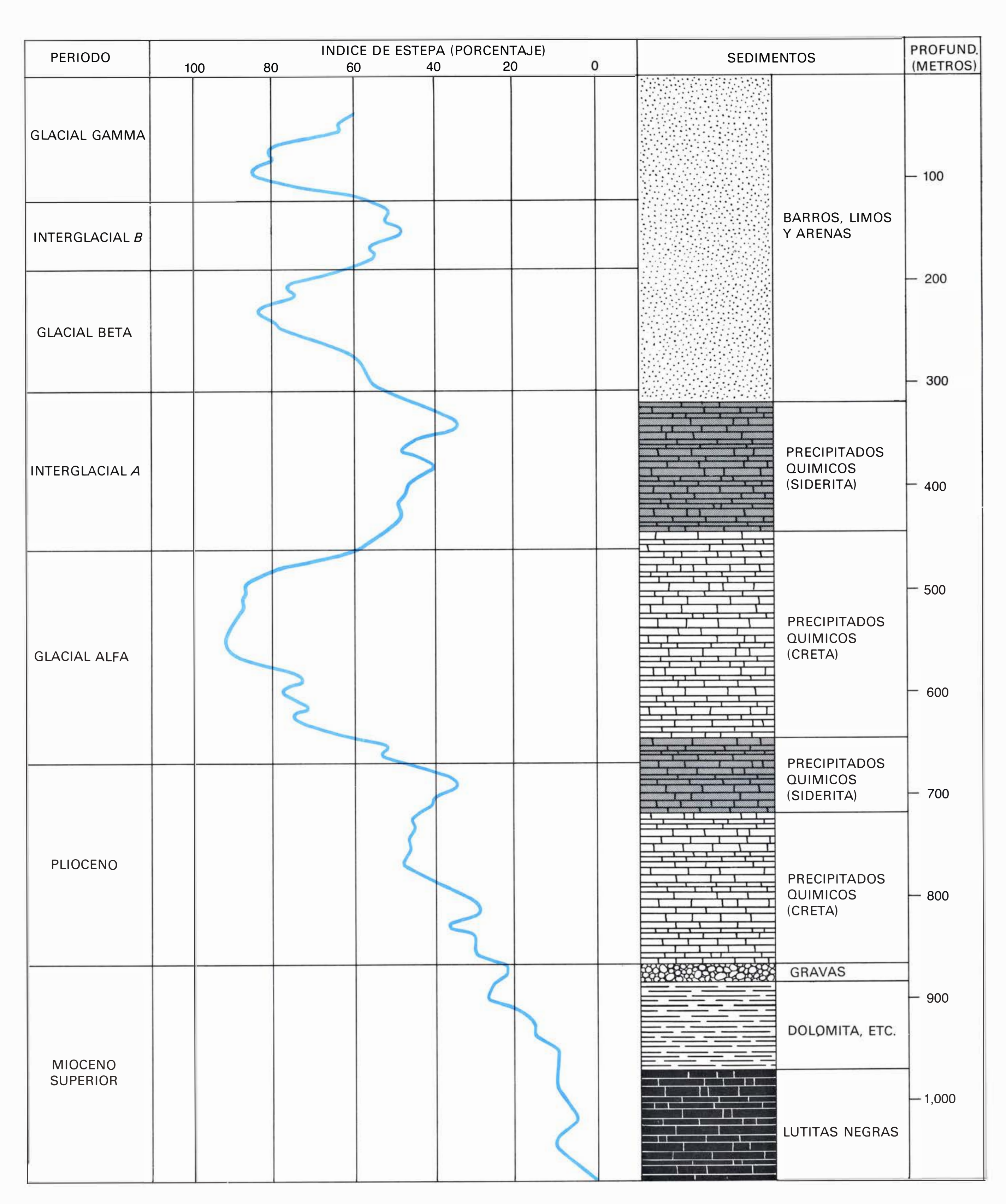
CIERRE DEL MAR TETIS, que se produjo por la colisión de Africa con Europa y Asia. Las posiciones aproximadas de los antiguos brazos de mar se superponen aquí sobre un mapa de las costas actuales. Hace unos 20 millones de años (*a*), el Tetis estaba aun abierto por el este y por el oeste, si bien se dividía en dos grandes brazos. El brazo meridional fue el antepasado del Mediterráneo actual, y al septentrional se le dio el nombre de Paratetis. La formación de montañas aisló finalmente el Paratetis (*b*) que se convirtió luego en un mar interior salobre. Hace unos seis millones de años, el continuado avance hacia el norte de la placa continental africana originó un episodio extraordinario en la his-

iniciarse la expedición, se había sospechado que el agua del mar Negro no había sido siempre salobre. En 1969 la Woods Hole Oceanographic Institution había obtenido testigos de sedimentos de poca profundidad (de unos 10 metros de longitud) en el mar Negro. Ross y Egon T. Degens, de la Woods Hole, encontraron pruebas en estos testigos de que el mar Negro fue un lago de agua dulce durante un período de 12.000 años en la última época glacial. Pudieron explicar el episodio de agua dulce como consecuencia de un cambio del nivel del mar producido a escala mundial. El Bósforo no solamente es un estrecho de poca anchura, sino que su profundidad también es pequeña, con un umbral situado a menos de 35 metros bajo el nivel actual del mar. Durante los períodos más fríos de la era glacial, cuando una gran parte del agua del mundo estaba almacenada en los glaciares continentales, el nivel del mar descendió hasta 100 metros por debajo de su nivel actual. Durante estos períodos, el Bósforo no era un estrecho sino un río meandriforme que drenaba un lago de agua dulce situado en la cuenca del mar Negro.

Se había supuesto que el episodio con agua dulce descubierto en 1969 era un incidente muy raro atribuible a condiciones climáticas poco comunes. Los sedimentos mucho más profundos recuperados por el *Glomar Challenger* demostraron que ello no era así. En todo caso, serían las condiciones marinas actuales del mar Negro las que debieran considerarse probablemente como ex-



toria del Mediterráneo (*c*). La formación de montañas cortó la comunicación con el Atlántico evaporándose completamente todo el agua y convirtiéndose en una cuenca desértica. La desecación catastrófica del Mediterráneo extendió su influjo, a la larga, en el Paretetis. Cuando el nivel del agua descendió, los ríos se rejuvenecieron y sus cabeceras se erosionaron abriéndose paso finalmente por la divisoria existente entre las cuencas. El resultado fue que el Paratetis desaguaba en el Mediterráneo y que ambos mares quedaron reducidos a un sistema de lagos (*d*). Al reorganizarse el sistema de drenaje, las cuencas del Paratetis, que no podían ser drenadas directamente, perdieron agua por evaporación.



REGISTRO SEDIMENTARIO del mar Negro obtenido en 1975 por el *Glomar Challenger,* buque de investigación diseñado expresamente para sondear el fondo marino. Este registro es del Emplazamiento 380, desde donde se extrajo un testigo con una longitud total de más de un kilómetro. Los sedimentos más profundos son lutitas negras formadas por compactación de fangos. Sobre esta capa hay sedimentos que contienen dolomita y gravas, indicadoras una y otras de un medio de aguas poco profundas. Vino después un largo período de sedimentación química durante el cual precipitaron alternativamente carbonato de calcio (creta) y carbonato de hierro (siderita). A 332 metros en profundidad, los sedimentos son predominantemente fangos, limos y arenas, que se siguen depositando hoy. Al principio no se podía datar esta serie sedimentaria, pero el polen fósil permitió la correlación con el clima. Un "índice de estepa" desarrollado por Alfred Traverse, de la Universidad estatal de Pennsylvania, indica el porcentaje de polen que procede de la vegetación esteparia característica de climas más fríos. En el índice de estepa se pueden reconocer tres períodos de glaciación. Estos episodios fríos vinieron precedidos por un período de condiciones semitropicales.

cepcionales, ya que durante una gran parte del período de tiempo representado en los sedimentos prevalecieron condiciones de agua dulce.

A consecuencia de este descubrimiento se ha visto que el mar Negro es un mal lugar para estudiar la génesis del petróleo. Excepto en un horizonte cercano a la superficie, ninguno de los sedimentos recuperados se mostraba especialmente rico en carbono orgánico. Las condiciones anóxicas que prevalecen actualmente en las aguas más profundas del mar Negro son, sobre todo, el resultado de la interacción existente en el Bósforo entre sus aguas salobres superficiales y las aguas más salinas del Mediterráneo. Hay un sistema de corrientes opuestas y estratificadas en el estrecho: el agua del mar Negro fluye por la superficie hacia el sur, mientras que la del Mediterráneo, que tiene una salinidad casi doble, avanza hacia el norte por el fondo. El agua más salina es más densa, por lo que se hunde en la cuenca, donde permanece confinada (debido al gradiente de densidad) y aislada del sistema de circulación del mar. Al no tener acceso a la superficie, el agua del fondo no puede disolver oxígeno y todas las formas de vida superior quedan absolutamente excluidas de la llanura abisal.

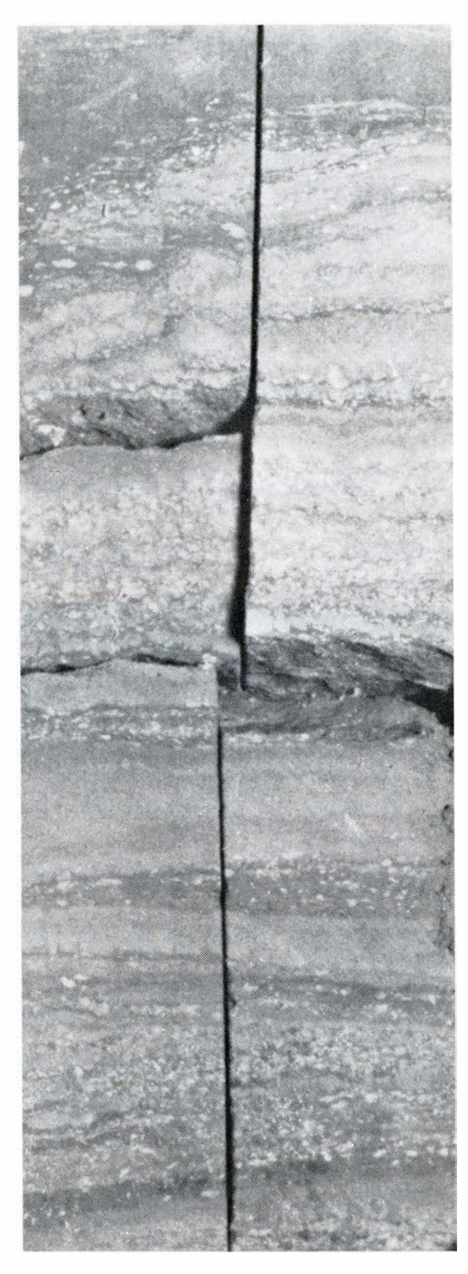
Este mecanismo depende de una entrada continua de agua marina y, ciertamente, no pudo haber actuado cuando el mar Negro era un lago de agua dulce. Las pruebas suministradas por fósiles de animales bentónicos encontrados en varios niveles de los sedimentos indica que las aguas del fondo fueron aireadas con bastante frecuencia en épocas pasadas. En esas condiciones, los detritus orgánicos que llegaban al fondo eran consumidos por los organismos que de ellos se alimentan. Otras veces, el fondo parecía hallarse desprovisto de vida, ya que los sedimentos, finamente laminados, no eran alterados por organismos bentónicos e, incluso entonces, se depositaba poca materia orgánica.

El predominio de sedimentos de agua dulce interponía también un obstáculo serio a la interpretación de la historia del mar Negro. El método normal de datar series sedimentarias consiste en buscar especies fósiles características de cada intervalo de tiempo geológico; en efecto, las divisiones convencionales del tiempo geológico se definen, sobre todo, por la aparición o desaparición de determinadas especies que se destacan en el registro fósil. Esas especies características son escogidas por su amplia distribución geográfica y por su limitada duración temporal. Casi todas ellas son marinas, por lo que resultó difícil datar de inmediato los sedimentos del mar Negro, en donde predominan los fósiles de agua dulce. Se desconocía incluso la duración total de la secuencia; el testigo de un kilómetro de longitud podían representar los sedimentos acumulados durante 10 millones de años, o los acumulados solamente durante dos o tres millones.

Los sedimentos del mar Negro contienen fósiles. Entre ellos se encuentran diatomeas y dinoflagelados planctónicos, foraminíferos bentónicos, determinados animales bentónicos de dimensiones mayores, como ostrácodos y moluscos, y polen y esporas de plantas terrestres que fueron a parar al mar. Casi todos los fósiles pertenecen a especies que no cambiaron durante largos períodos y son, por tanto, de poca utilidad para datar los sedimentos. Sí prestan, en cambio, abundante información en lo referente a los cambios de condiciones en el mar Negro, incluyéndose entre ellos los correspondientes al clima de las regiones circundantes.

En las plantas los fenómenos de especiación y cambio evolutivo son generalmente procesos bastante lentos, pero su desarrollo es muy sensible al clima. Por tanto, el material fósil procedente de plantas es un buen indicador de la temperatura de épocas pasadas. En los sedimentos del mar Negro se encuentran fósiles pertenecientes a decenas de géneros de plantas, principalmente en forma de granos de polen, que son característicos e inmediatamente identificables. Las plantas presentan unos hábitat amplios, que pueden agruparse en dos categorías generales. En algunos estratos el polen procede principalmente de arbustos específicos de la vegetación esteparia, mientras que en otros proviene de ciertos árboles, verbigracia, pinos, robles y hayas, lo cual indica la presencia de bosques. Alfred Traverse, de la Universidad estatal de Pennsylvania, que estuvo en la expedición al mar Negro, ideó un "índice de estepa" para expresar cuantitativamente la naturaleza de la flora terrestre. Este índice es simplemente el porcentaje de polen encontrado en un nivel determinado que procede de vegetación esteparia. Un valor de 100 implica que la cuenca hidrográfica del mar Negro estuvo cubierta casi enteramente por estepas y un valor de cero significa un paisaje con bosques constituidos por una mezcla de pinos y árboles de hoja caduca.

En la mayor parte del registro sedimentario alternan la vegetación esteparia y forestal. Esta alternancia se interpreta fácilmente en función del clima. Los bosques predominaron durante los períodos interglaciales cálidos, pero con el avance de los glaciares continentales fueron reemplazados por estepas. La curva que nos marca el índice de estepa tiene muchas fluctuaciones de período corto, mas al aplicar un procedimiento pro-



DOLOMITA LAMINADA, procedente de un testigo profundo extraído en el Emplazamiento 380. Sirve para demostrar que el mar Negro fue anteriormente un lago somero. En la dolomita se encuentran unas estructuras denominadas estromatolitos (del griego, "piedra plana"). Los estromatolitos se forman en virtud del desarrollo de algas azulverdosas que tejen sobre el fondo delgadas mallas. Puesto que las algas son fotosintéticas sólo pueden desarrollarse en aguas poco profundas, habitando hoy en día sobre todo en plataformas de marea. Las gravas situadas encima de los estromatolitos son también sedimentos típicos de aguas someras.

mediador que suaviza la curva, Traverse pudo identificar tres períodos glaciales principales que designó por Alfa, Beta y Gamma.

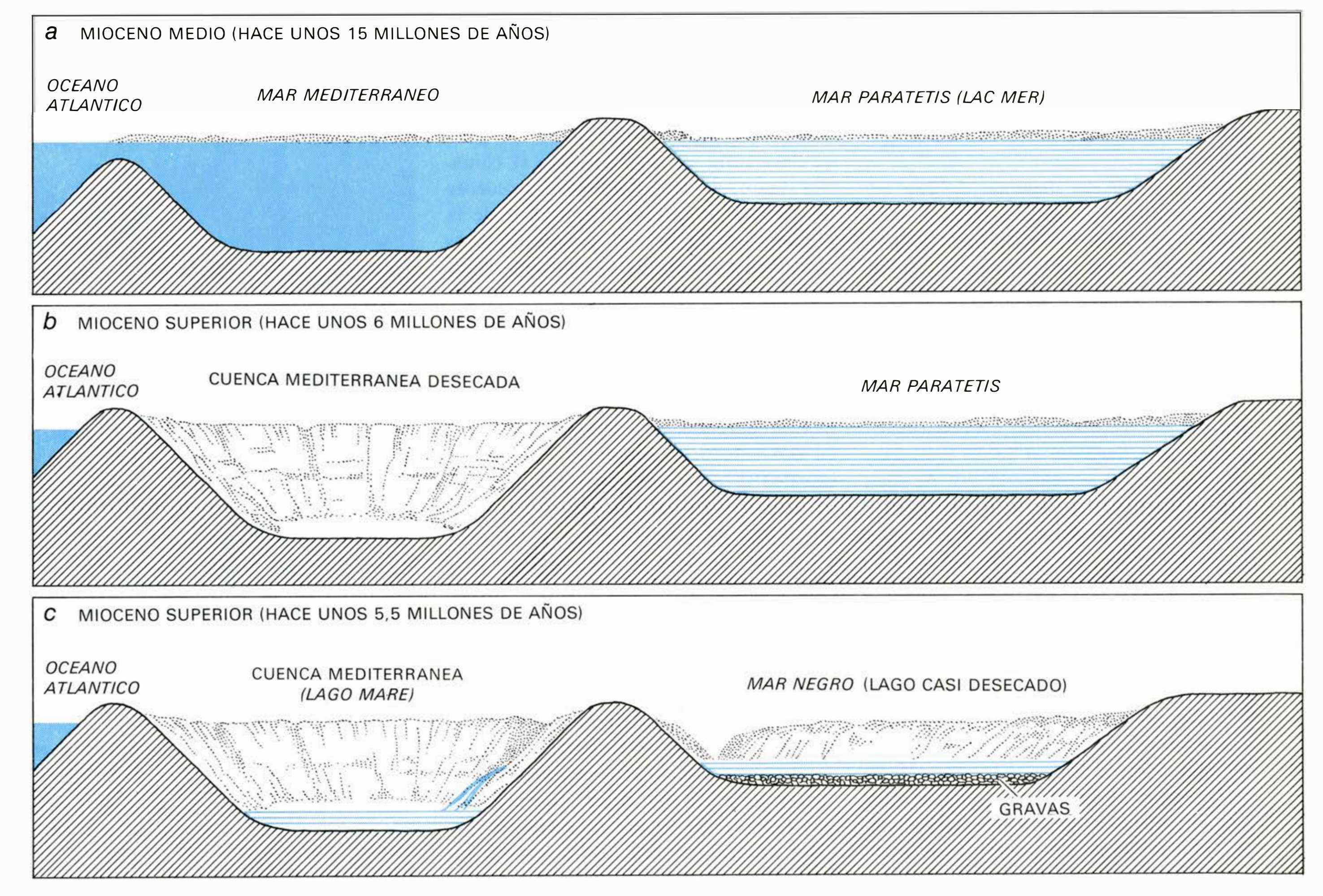
Las oscilaciones climáticas sugeridas por los análisis de polen no se continúan en los sedimentos más antiguos obtenidos en los sondeos. En el testigo del Emplazamiento 380 el polen de vegetación esteparia es raro o falta a profundidades mayores de unos 650 metros por debajo del fondo marino. En los estratos más inferiores, el polen procede principalmente de plantas semitropicales. Así pues, resultaría que el registro sedimentario se extiende hasta una época anterior al inicio de las glaciaciones. Por desgracia, las muestras de polen no aportan información alguna sobre la fecha del comienzo o sobre las fechas de los tres períodos de glaciación subsiguientes. Con ese desconocimiento de la cronología, todas las series sedimentarias permanecían como una especie de texto inquietante que no podía leerse. Y fue así hasta hace dos años, cuando un equipo internacional de investigadores encontró la clave de su mensaje.

Rasgo sobresaliente de los testigos del mar Negro es la transición abrupta que se da entre dos tipos completamente distintos de sedimentos. En un largo intervalo, el material depositado en el fondo era principalmente de precipitados químicos; luego, en un momento dado, definido por un límite situado a 332 metros por debajo del fondo marino, el tipo de sedimento dominante pasaba a estar constituido por limos y fangos, que aún se continúan formando en la actualidad.

Los sedimentos químicos son minerales carbonatados: calcita ($CaCO_3$), aragonito ($CaCO_3$), dolomita ($CaMg(CO_3)_2$) y siderita ($FeCO_3$). El que un lago de agua dulce pudiera estar saturado de carbonatos no es nada sorprendente. La creta, constituida casi exclusivamente por calcita, precipita cada verano sobre el fondo del lago suizo de Zurich.

Con el aumento de temperatura el anhídrido carbónico se escapa de la solución, lo que a su vez provoca un descenso en la solubilidad de los minerales carbonatados, en virtud de lo cual precipitan. La existencia de un lago en la cuenca del mar Negro durante las glaciaciones hubiera implicado un medio similar al que tiene el lago Zurich en la actualidad: los glaciares se encontraban a una cierta distancia y la variación anual de la temperatura era suficiente para alterar la solubilidad de los carbonatos. La formación de carbonatos distintos de la calcita puede explicarse por cambios en la química del agua del lago. El carbonato de hierro, denominado siderita, por ejemplo, precipitó sobre todo en épocas cálidas y húmedas cuando los ríos divagantes llevaban mucho hierro disuelto hasta la cuenca lacustre.

Los minerales puros sólo pueden



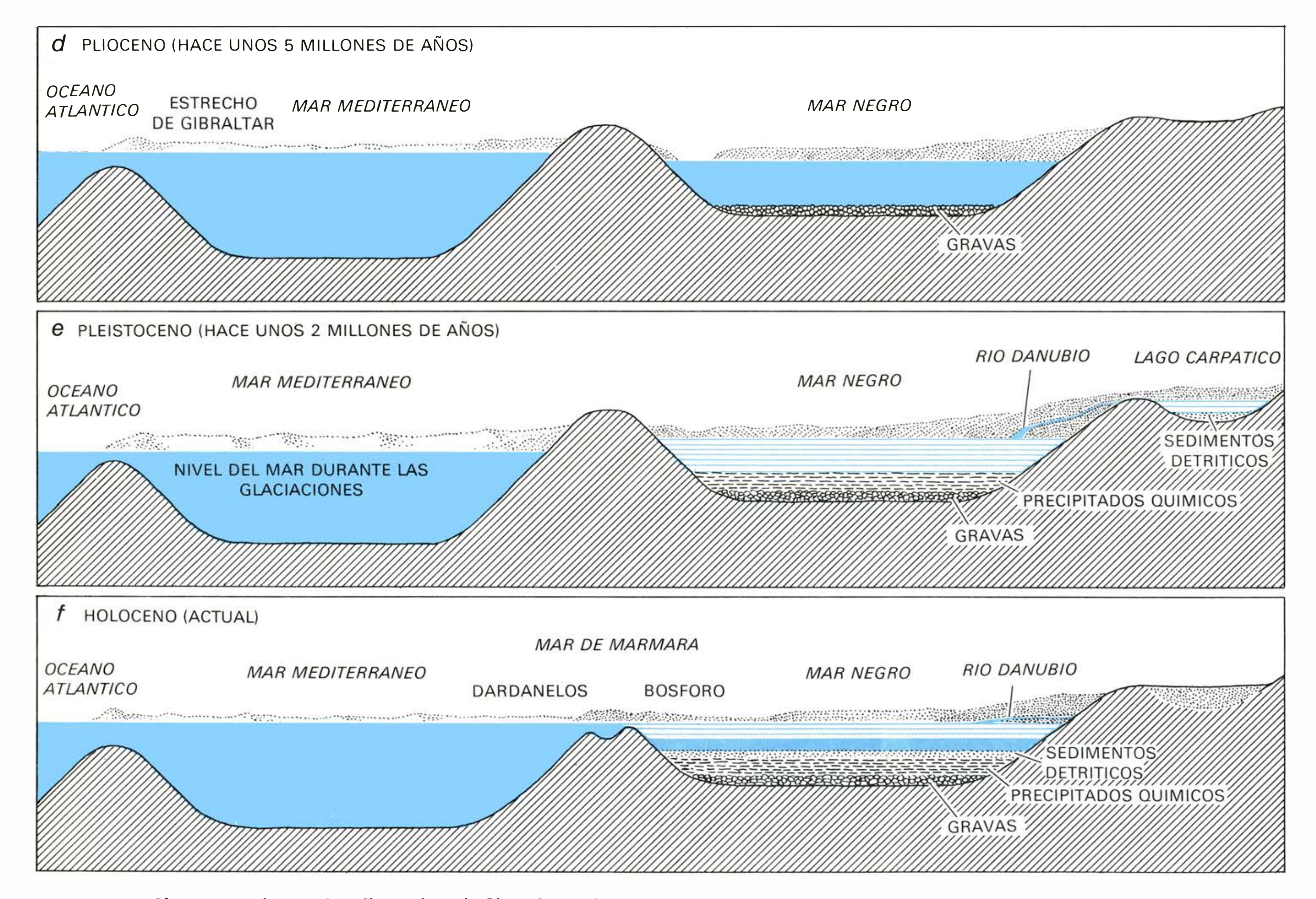
HISTORIA AFIN del Mediterráneo y del mar Negro, presentada esquemáticamente en un corte de las dos cuencas. Cuando los dos mares se separaron, debido al levantamiento de diversas cordilleras alpinas (*a*), el Paratetis recibía una importante escorrentía, procedente de las llanuras de Europa Central, que lo mantenía lleno. Sin embargo, su salinidad disminuía. Durante la subsiguiente desecación del Mediterráneo (*b*), el Paratetis se mantuvo en estas condiciones hasta que la reorganización del sistema de drenaje europeo desvió gran parte de su agua hacia el Mediterráneo. Durante esa época (*c*) fue cuando se depositaron los sedimentos de aguas poco profundas, como las gravas. Cuando el estrecho de Gibraltar funcionó de nuevo, llenando otra vez el Mediterráneo, en el mar Negro (y posiblemente otras partes del Paratetis)

precipitar en un lago si las arcillas, limos y los otros productos resultantes de la meteorización del paisaje circundante (que reciben el nombre de clásticos o sedimentos detríticos) no se depositan en él. En lo que se refiere al lago Zurich los clásticos se sedimentan en otros lagos, aguas arriba. Por otro lado, el mar Negro recibe en la actualidad gran cantidad de clásticos que acarrea, sobre todo, el Danubio. En tiempos anteriores debió existir alguna barrera para evitar que los clásticos del Danubio llegaran a la cuenca del mar Negro.

La naturaleza de esta barrera quedó aclarada en 1976, cuando el geólogo rumano Dan Jipa me mostró un potente corte de materiales detríticos al pie de los Cárpatos orientales. El material se encuentra hoy al aire libre, pero fue depositado en forma de sedimento en un lago de edad pliocena y pleistocena. Durante una visita realizada al delta del Danubio se me indicó que los sondeos allí realizados llegaban al substrato rocoso después de atravesar un delgado recubrimiento de sedimentos del Pleistoceno. Está claro que el Danubio pasó alguna vez por el lago de los Cárpatos donde una gran parte de los sólidos que aquél llevaba en suspensión tenía tiempo de sedimentarse. En un momento determinado del Pleistoceno el lago se rellenó de limo, por lo que el Danubio se vio obligado a cambiar su curso. Privado de este "tanque de sedimentación" situado aguas arriba, el mar Negro quedaría invadido por los materiales detríticos llegándose así al final del período de sedimentación química.

Los sedimentos de los Cárpatos proporcionaron la primera información sobre la edad de los sedimentos del mar Negro, si bien la datación de los mismos era muy aproximada. Se estimó que el techo de las series de los Cárpatos tenía una edad comprendida entre 0,5 y 1,5 millones de años. Si la transición de la creta a los fangos, situada a una profundidad de 332 metros, corresponde a un cambio en el curso del Danubio, entonces este horizonte debe haberse depositado durante el mismo intervalo.

Una datación más precisa de los sedimentos pleistocenos del mar Negro se obtuvo, por fin, al estudiar el magnetismo remanente de las rocas sedimentarias. Muchos minerales, en particular los que contienen hierro, son débilmente magnéticos y sus campos tienden a alinearse con el campo magnético que la tierra posee cuando se depositan. El campo de la tierra ha invertido ocasionalmente su polaridad; y se ha catalogado y datado estos acontecimientos basándose en amplios estudios de las rocas sedimentarias del fondo del Pacífico y del Atlántico. Si las inversiones de la orientación magnética se podían encontrar en los testigos del mar Negro, era posible entonces datar estos últimos



se reanudó por poco tiempo el medio marino (*d*). Sin embargo, la comunicación con el Mediterráneo se cortaría de nuevo muy pronto, y el mar Negro empezó una larga etapa como lago de agua dulce (*e*). Los limos y otros productos de meteorización, que conjuntamente reciben el nombre de sedimentos detríticos, quedaron represados en un lago de los Cárpatos orientales, de modo que los minerales carbonatados pudieron precipitar en el mar Negro. Hace unos 600.000 años, el lago de los Cárpatos fue colmatado y los sedimentos detríticos, arrastrados principalmente por el Danubio, llegaron al mar Negro. En esta época se inicia la formación de depósitos de limos y fangos. Más tarde, hace unos 100.000 años, el Bósforo se vio invadido por las aguas marinas que de nuevo entraron en el mar Negro (*f*), produciendo un estrato anóxico de aguas salobres.

refiriéndolos a la cronología encontrada en otras partes.

Resultó difícil detectar las inversiones del campo, porque muchos minerales de los sedimentos del mar Negro que contienen hierro no son estables químicamente. La orientación magnética que presentan hoy podría reflejar la dirección del campo que la tierra tenía, no cuando los minerales precipitaron por primera vez, sino en el tiempo correspondiente al de alguna transformación química ocurrida posteriormente. Sin embargo, mi adjunto Federico Giovanoli y yo mismo pudimos identificar dos inversiones que creemos corresponden a los acontecimientos denominados Jaramillo y Olduvai en las cronologías compiladas en otras partes del mundo. Representan dos episodios breves, que ocurrieron hace 1 millón y 1,7 millones de años respectivamente, durante los cuales el campo de la tierra volvió a tener temporalmente una polaridad normal que se intercaló en una prolongada época en que el campo estuvo invertido de forma predominante.

El descubrimiento de las anomalías magnéticas proporcionó dos dataciones fijas en la historia de la variación climática puesta de manifiesto por los estudios de polen. Nuestra interpretación de la cronología sugiere que la glaciación continental empezó, en las regiones que bordean el mar Negro, a principios del Pleistoceno, hace 1,8 millones de años, o en época cercana. Los datos revelan igualmente que existieron tres períodos principales de glaciación junto con un gran número de períodos cortos de clima frío. La correlación de los períodos del mar Negro designados Alfa, Beta y Gamma con los conocidos de otras regiones no es segura. Se hace tarea punto menos que imposible hallar alguna correlación con los períodos glaciares clásicos de Europa. Sin embargo, los tres períodos principales y los numerosos episodios cortos de clima frío parecen corresponderse con las variaciones climáticas registradas en los sedimentos marinos del Atlántico. Así, por ejemplo, existe una correlación muy estrecha con un registro climático reconstruido por M. Briskin y William A. Berggren, de la Woods Hole Oceanographic Institution, utilizando fósiles de los sedimentos atlánticos. La interpretación de las inversiones magnéticas proporcionó igualmente una edad para el límite situado a 332 metros, que indica la primera llegada al mar Negro de los sedimentos detríticos del Danubio. El cambio del curso de este río parece haber ocurrido hace unos 600.000 años.

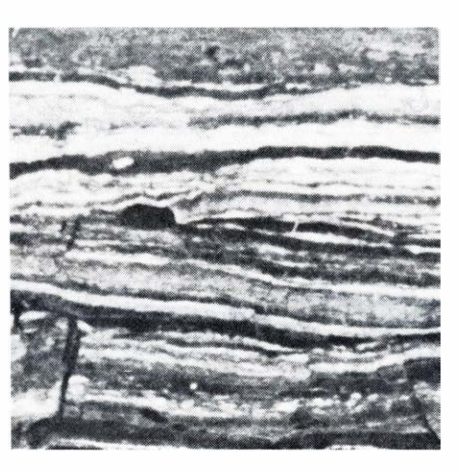
Los estudios del polen fósil y del magnetismo remanente proporcionaron finalmente la cronología del mar Negro durante todo el Pleistoceno, es decir, de aproximadamente los dos últimos millones de años. Una primera incursión hacia épocas más antiguas se debió a Musat Gheorghian, de la Academia Rumana de Ciencias, quien encontró en las muestras más antiguas procedentes de los testigos del mar Negro una fauna fósil típica del Paratetis. Los fósiles pertenecían a diversas especies relictas de foraminíferos bentónicos, propios de un mar salobre, y su desarrollo evolutivo sugirió que los sedimentos más antiguos recuperados del Emplazamiento 380 tenían una edad comprendida entre ocho y diez millones de años. Las pruebas paleobotánicas apoyaron esta conclusión: en las mismas capas de mayor antigüedad Traverse había encontrado polen abundante, característico de la vida vegetal propia de un hábitat cálido y de tierras altas. Las plantas en cuestión desaparecieron de Europa hace entre seis y ocho millones de años aproximadamente, cuando el clima empezó a enfriarse apreciablemente. Si esta datación de los sedimentos es correcta, el registro debería extenderse entonces por el Mioceno superior, cuya edad está comprendida entre cinco y seis millones de años, durante el cual el Mediterráneo se secó. ¿Qué ocurrió en el mar Negro cuando su vecino quedó reducido a desierto?

Antes de la expedición de 1975 del *Glomar Challenger*, una expedición oceanográfica francesa al mar Negro había encontrado un horizonte fuertemente reflector de ondas acústicas situado a 1000 metros, o más, por debajo del fondo marino. También se hallaron estructuras profundas fosilizadas, parecidas a las formaciones salinas. Naturalmente, el descubrimiento de un reflector de este tipo fue lo que proporcionó el primer indicio para desarrollar la hipótesis de la desecación del Mediterráneo. El *Glomar Challenger* perforó efectivamente un horizonte reflector, pero no se sabe aún con certeza si es el mismo que encontró la expedición francesa. En todo caso no se halló sal en los sedimentos, y el reflector cortado a 865 metros de profundidad en el Emplazamiento 380 consistía en gravas cementadas.

Las gravas son sedimentos de escasa profundidad. Ocasionalmente se transportan desde una zona costera hasta una plataforma profunda por corrientes de turbidez submarinas, pero las del mar Negro no presentaban señal alguna que recordara ese transporte; y así fue tomando cuerpo la idea de que, cuando las gravas se sedimentaron, el mar Negro era un lago somero enclavado en una cuenca profunda.

El examen de los fósiles asociados a las gravas cementadas y a las capas adyacentes suministró pronto argumentos para apoyar esta hipótesis. Peter Stoffers, de la Universidad de Heidelberg, encontró dolomita entre los niveles de grava. En la dolomita aparecían unas estructuras denominadas estromatolitos, formados por el desarrollo de algas azulverdosas, constituyendo las denominadas llanuras de mallas de algas ("flat mats"), que dan a la roca una estructura laminar característica. Fondos con crecimiento de algas similares se dan en regiones costeras actuales y, puesto que estas plantas necesitan la luz del

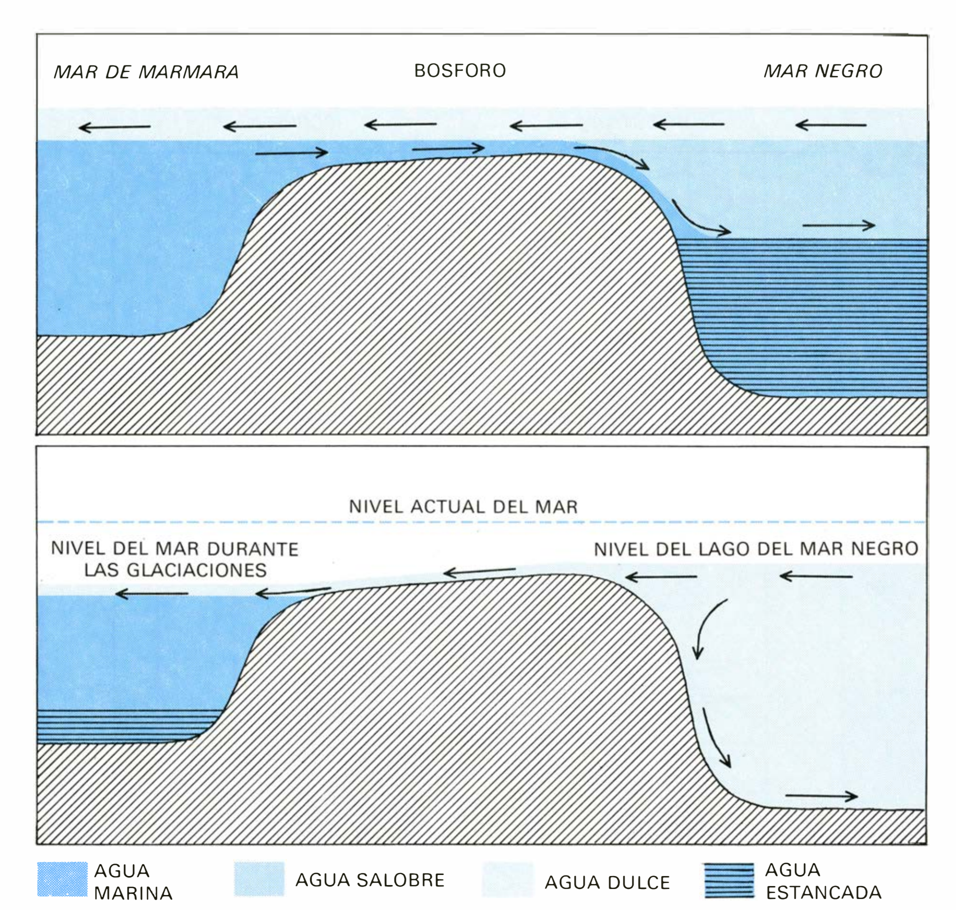




DEPOSICIONES DE CRETA del Emplazamiento 380 que nos demuestran de la existencia de períodos alternantes. En unos, las profundas aguas del mar Negro eran dulces y estaban habitadas, mientras que en los otros eran salobres y prácticamente sin vida. Las marcas irregulares y obscuras de los sedimentos de la izquierda son huellas de pequeños gusanos excavadores bentónicos. Los sedimentos de la derecha están constituidos por capas finas y alternantes de creta y material detrítico arrastrado hasta el mar por los ríos. Si hubiera existido algún tipo de animales bentónicos cuando se depositaron estos sedimentos, la estructura finamente laminada se hubiera alterado.

sol para realizar la fotosíntesis, sólo pueden desarrollarse en aguas poco profundas. Stoffers halló también una roca formada esencialmente por oolitos, los cuales constituyen un sedimento arenoso típico de los canales de marea próximos a la línea de costa. Hans Schrader, de la Universidad de Kiel, identificó en sedimentos encontrados precisamente debajo de las gravas una especie de diatomea característica de un hábitat de aguas poco profundas.

Al haber yo compartido las responsabilidades de director científico de las dos expediciones llevadas a cabo en el Mediterráneo con objeto de efectuar perforaciones, me entusiasmaba la posibilidad de demostrar la existencia de un episodio de desecación en el mar Negro, perteneciente al Mioceno superior. Sin embargo, había que dar con pruebas concluyentes. En primer lugar, ninguno de los fósiles encontrados en las gravas y en los sedimentos asociados podía datarse con facilidad. Además, existían pocos argumentos que demostrasen que el mar Negro pudo haberse desecado, lo cual resultaba aún más preocupante. A diferencia del Mediterráneo, el mar Negro que hoy contemplamos no tiene déficit hidrológico alguno: los aportes de varios ríos principales compensan de sobras las pérdidas por evaporación. Por otra parte, el mar Negro fue una de las cuencas más profundas del Paratetis. Si hubiera perdido la mayor parte de su agua, se hubiera desecado todo el mar. ¿Habría alguna prueba que explicase esta desecación generalizada?

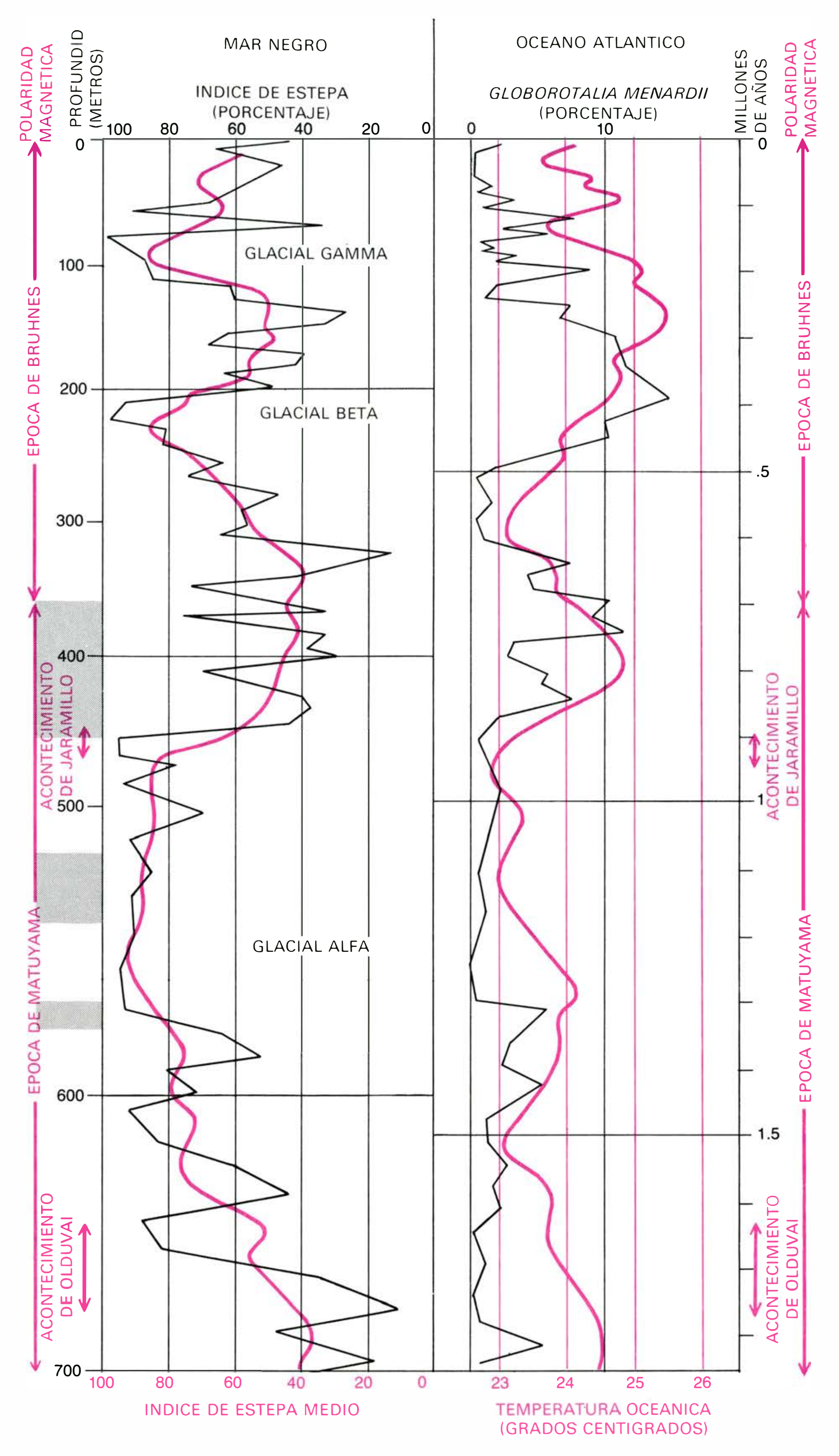


LAS CORRIENTES DEL BOSFORO han tenido una influencia capital en el medio físico del mar Negro. Hoy en día *(arriba)*, el umbral del Bósforo está unos 35 metros por debajo el nivel del mar y existe un sistema de contracorrientes estratificadas. Del mar Negro sale agua comparativamente dulce por la superficie, al mismo tiempo que entra por el fondo agua marina más salada. Puesto que el agua salina es más densa, se hunde y queda estancada. Durante las glaciaciones, el nivel general de los mares bajó hasta 100 metros por debajo del actual. A lo largo de tales períodos glaciares *(parte inferior)* el Bósforo no era el estrecho que hoy conocemos, sino un río, y el lago de agua dulce existente en la cuenca del mar Negro quedaba aireado a cualquier profundidad.

En 1976, mientras me hallaba realizando un viaje de intercambio científico en Europa oriental, empezó a perfilarse un modelo coherente. Existen sedimentos en regiones de la cuenca del Paratetis que no estaban siempre sumergidas y que contienen capas de cenizas volcánicas, las cuales pueden datarse midiendo el grado de desintegración de los isótopos radiactivos desde que el material pasó por el estado de fusión. Gracias a este método se estableció que los niveles salinos encontrados en determinadas partes de la cuenca del Paratetis tienen una edad superior a los 15 millones de años, muy superior a la de los niveles salinos del Mediterráneo. Durante el Mioceno superior no hubo precipitación de sales en lo que hoy en día se denominan países balcánicos y, por tanto, es improbable que el Paratetis se secara por evaporación. Por otra parte, R. Jiříček, de la Academia Checoslovaca de Ciencias, pudo demostrar que, en realidad, el Paratetis de los Balcanes sufrió una crisis cuando el Mediterráneo se secó. En esta época el salobre Lac Mer se desintegró dejando solamente una serie de lagos de agua dulce diseminados. Puesto que no precipitó sal, las cuencas lacustres no pudieron evaporarse hasta la sequedad, y la única alternativa verosímil era que una gran parte del agua del Paratetis desapareciera gracias a un sistema de drenaje.

El trabajo de Jiříček proporcionó la información necesaria para reconstruir la historia del Mediterráneo y la del mar Negro, tan complicadas y al propio tiempo estrechamente relacionadas. Durante el Mioceno superior, el Mediterráneo era una llanura salina situada en el fondo de una profunda depresión, como si se tratara de un enorme Valle de la Muerte. Posteriormente quedó invadido por agua salobre, creando una serie de lagos en el fondo de la cuenca. A este sistema lacustre se le ha dado el nombre de Lago Mare, equivalente italiano de Lac Mer. En estos lagos vivieron animales de aguas salobres. Entre ellos se encontraban los ostrácodos del género *Cyprideis* cuyos fósiles, hallados encima de las formaciones salinas, pero situados por debajo los sedimentos marinos depositados después de la apertura del estrecho de Gibraltar, habían constituido un verdadero enigma durante la investigación del Mediterráneo.

El agua necesaria para llenar el Lago Mare provenía del Paratetis, el cual, debido a ello, era vaciado parcialmente. Con la desecación del Mediterráneo, el consiguiente rejuvenecimiento de los cursos fluviales había entallado profundos cañones en el talud continental y, si bien esas zonas están de nuevo hoy en día sumergidas, algunos de estos cañones pueden observarse aún en algunos perfiles del fondo marino. Al mismo tiempo, la erosión producida en las cabeceras aumentó la longitud de los ríos, haciendo que la divisoria de aguas entre la cuenca mediterránea y el Paratetis se desplazara hacia el norte. Por último, esta divisoria se cortó en algún sitio y se vació el Lac Mer debido



CRONOLOGIA de los sedimentos del mar Negro. Se pudo confeccionar a partir de la identificación de las inversiones de la orientación magnética en las rocas procedentes de los testigos de sondeo. Las anomalías magnéticas se produjeron por inversiones de la polaridad del campo magnético de la tierra, que pueden detectarse igualmente en otras rocas sedimentarias, entre las que se encuentran algunas que están bien datadas. Durante los acontecimientos Jaramillo y Olduvai, y durante la época de Bruhnes, el campo tenía la misma orientación que tiene hoy en día, mientras que durante la época de Matuyama se hallaba invertido. En algunos niveles sedimentarios *(gris)* la dirección del campo no está precisada suficientemente. El registro climático del mar Negro, representado por el índice de estepa, se puede comparar como mínimo con dos fechas, fijadas por las anomalías magnéticas y con otros indicadores del clima antiguo. Uno de estos indicadores, que se correlaciona con la temperatura media del océano, es el porcentaje existente en los sedimentos marinos del foraminífero *Globorotalia menardii.* Los valores medidos de ambos indicadores pueden observarse en las curvas en negro y en color. La curva en color de la derecha da una estimación de las temperaturas basadas en varias especies fósiles, además de la del foraminífero *G. menardii.*

probablemente, a los profundos cursos fluviales que desde Hungría y Yugoslavia desaguaban en el Adriático.

Con la formación del Lago Mare se reorganizó el sistema hidrográfico de Europa. El Paratetis había estado recibiendo gran parte de las aguas de las regiones húmedas de Europa central, mientras que ahora era la cuenca mediterránea quien capturaba su flujo. Debido a ello el mar Negro ya no recibía agua suficiente para compensar la evaporación. Puesto que no se encontraron los horizontes salinos en los testigos de los sondeos, no puede determinarse con certeza si el mar Negro resecó alguna vez completamente, pero es indudable que durante algún tiempo fue un lago de salinidad elevada y de poca profundidad. Se depositaron entonces gravas y fósiles pertenecientes a especies de aguas someras; además, Frank T. Manheim, de la Southern Florida State University, al estudiar el agua existente en los intersticios de los sedimentos ha hallado que este lago de profundidad escasa tenía como mínimo una salinidad tres veces superior a la del agua marina normal.

El drenaje del Paratetis hubiera transportado especies animales típicas de este mar hacia el Mediterráneo. Estos animales pueden muy bien haber sido los antepasados de las especies relictas y aisladas que se encuentran en la actualidad en las grutas de la costa yugoslava y en las regiones costeras de otras zonas del Mediterráneo. Durante el Mioceno ocuparon el sistema de lagos del Lago Mare; cuando el estrecho de Gibraltar se abrió, hubo una invasión de fauna marina, quedando exterminadas la mayoría de las especies residentes. Sólo las que se retiraron hacia hábitat de agua dulce aislados sobrevivieron. Los “fósiles vivientes” del lago Ohrid, en la frontera entre Yugoslavia y Albania, pueden tener un origen parecido; el lago es probablemente un resto de un antiguo brazo de mar que conectaba el Paratetis y el Mediterráneo, habiendo podido quedar abandonadas allí estas especies cuando se cortó la comunicación.

Las pruebas necesarias para confirmar esta hipótesis debían consistir en señales directas de que las gravas del mar Negro se depositaron durante el período en que el Mediterráneo se desecó. Esto se obtuvo a finales de 1976 cuando Schrader descubrió una flora de diatomeas, en unos sedimentos situados algo por debajo de las gravas, que podían datarse como pertenecientes al Mioceno superior.

El Lago Mare tuvo una vida corta. Se inundó cuando el agua de mar invadió el Mediterráneo a finales del Mioceno o principios del Plioceno hace unos 5,2 millones de años. En el mar Negro los fangos marinos o de aguas salobres situados directamente encima de los niveles de grava indican que la influencia marina se extendió brevemente, como mínimo hasta determinadas partes de la cuenca del Paratetis. El agua marina pudo haber entrado a través de los mismo canales que habían hecho desaguar anteriormente el Paratetis dentro de la cuenca mediterránea, si bien existen otras posibilidades. Cualquiera que sea el camino, la débil comunicación con el Mediterráneo quedó pronto cortada, el sistema de drenaje fue reorganizado de nuevo y el mar Negro empezó otra vez a recibir más agua de la que necesitaba para equilibrar sus pérdidas por evaporación. La cuenca del mar Negro se llenó de nuevo, pero el conjunto del Paratetis nunca se recuperó y la mayor parte de la extensión que otrora fue el fondo de este antiguo mar está en la actualidad intensamente cultivada y dedicada a la agricultura.

Con un exceso de aporte de agua continental el mar Negro perdió gradualmente su salinidad, y durante la mayor parte de los siguientes cinco millones de años fue un lago de agua dulce. A principios del Pleistoceno, hace 1,8 millones de años, el clima se volvió especialmente frío, formándose glaciares en Europa oriental. En esta época dominó la sedimentación química, y los sedimentos detríticos arrastrados por el Danubio se depositaban en un lago de los Cárpatos orientales. Hace algo más de medio millón de años el lago se colmató, el Danubio discurrió por su curso actual y los fangos se convirtieron en los sedimentos dominantes del mar Negro.

Hubo que esperar a épocas recientes para que el mar adquiriese sus características actuales, gracias a la erosión que fue abriendo gradualmente el Bósforo. Durante los episodios cálidos, cuando el nivel general de los mares se volvía excepcionalmente alto, las aguas marinas inundaban el Bósforo y, a través suyo, convertían en aguas salobres las del mar Negro; pero cada vez que los glaciares se instalaban de nuevo, se restablecía el régimen de agua dulce. Finalmente, a principios del Holoceno, hace aproximadamente 10.000 años, los océanos ascendieron hasta su nivel actual y el agua salada entró de nuevo por el Bósforo. Así nacieron las zonas abisales sin vida del mar Negro actual.

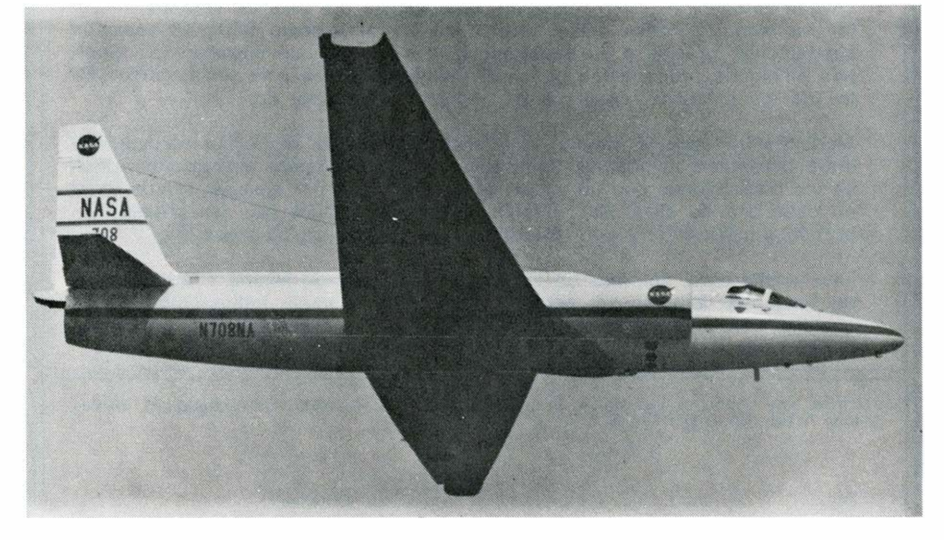
La radiación cósmica de fondo y el nuevo movimiento del éter

Instrumentos sensibles han detectado ligeras desviaciones de la uniformidad en la radiación originada por la primitiva "gran explosión". El experimento revela el movimiento de la Tierra con respecto al conjunto del Universo

Richard A. Muller

Una curiosa radiación que envuelve la Tierra casi uniformemente desde cualquier dirección ha resultado ser una fuente única de información sobre la naturaleza e historia del Universo. Esa débil radiación se descubrió hace 13 años durante una búsqueda de fuentes de ruido capaces de interferir con los sistemas de comunicación de satélites. Se demostró que el "ruido" era de origen cósmico y pronto se lo conoció como radiación cósmica de cuerpo negro a tres grados, porque tiene las características espectrales de un cuerpo negro, o emisor perfecto de radiación, cuya temperatura fuese de unos tres grados Kelvin (tres grados Celsius por encima del cero absoluto). La mayor parte de los astrofísicos creen ahora que esta radiación en microondas fue emitida inmediatamente después de la "gran explosión", o explosión cataclísmica en la que se creó el Universo hace unos quince mil millones de años. No sólo es la señal más antigua que nunca se haya detectado; es también la más distante, puesto que viene desde bastante más allá de los quasars, que son las más remotas fuentes luminosas conocidas. La radiación de tres grados es un telón de fondo delante del cual aparecen todos los objetos astrofísicos.

La observación de la radiación cósmica de fondo es lo más parecido al estudio directo de la explosión primaria misma. La propia existencia de la radiación es el argumento más sólido en favor de la teoría de la gran explosión. La isotropía de esta radiación, es decir, su uniformidad en diferentes direcciones en el espacio, nos dice que la gran explosión, aunque inimaginablemente violenta, se produjo con gran suavidad. La ligera desviación respecto a la isotropía que recientemente se ha descubierto indica que nuestra galaxia está atravesando el Universo a la sorprendentemente alta velocidad de 600 kilómetros por segundo. Esta velocidad cosmológica es lo que se ha llamado "el nuevo movimiento del éter", refiriéndose al "movimiento del éter" que A. A. Michelson y E. W. Morley trataron de descubrir, sin éxito, hace casi cien años midiendo la velocidad de la luz a lo largo de direcciones que formaban distintos ángulos respecto al movimiento de la Tierra en el espacio. La radiación cósmica de fondo de tres grados suministra un "éter" de radiación omnipresente para realizar un experimento análogo.



LA PLATAFORMA PARA INSTRUMENTOS en el experimento del nuevo movimiento del éter era un avión U-2 al servicio de la Administración Nacional de Aeronáutica y del Espacio (NASA). Como en el experimento original del movimiento del éter, realizado hace casi un siglo por A. A. Michelson y E. W. Morley, el nuevo experimento se proyectó con el fin de medir el movimiento de la Tierra con respecto a un sistema universal de referencia, en este caso la radiación cósmica de fondo. Tal radiación, que es equivalente a la emitida por un cuerpo negro (un radiador perfecto) cuya temperatura sea de unos tres grados Kelvin, es la radiación residual del cataclismo en que se creó el Universo hace 15.000 millones de años. El U-2 ha hecho diez vuelos llevando un receptor ultrasensible de microondas proyectado por el autor, George F. Smoot y Marc. V. Gorenstein.

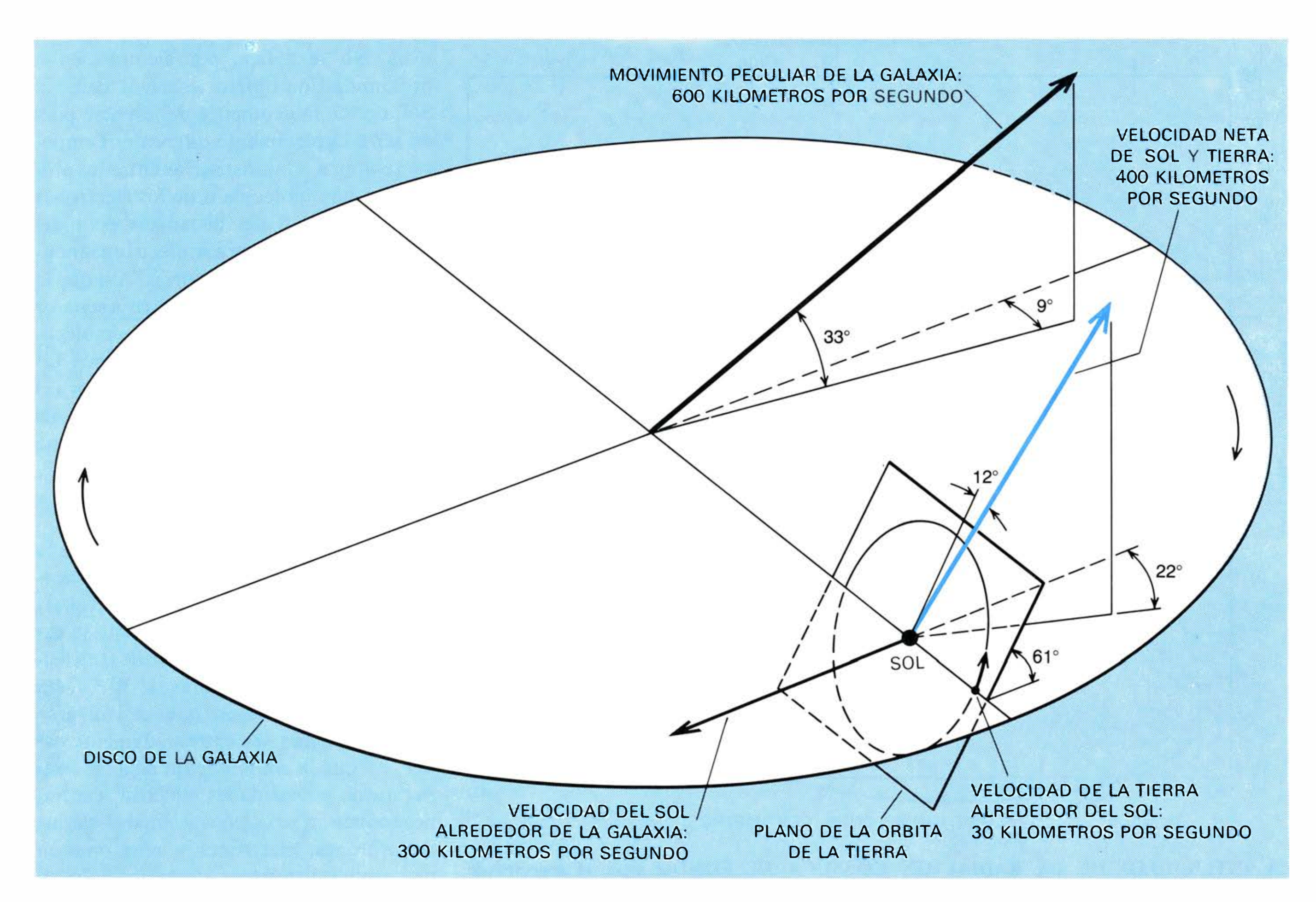
La radiación cósmica de fondo fue descubierta en 1965 por Arno A. Penzias y Robert W. Wilson, de los Laboratorios Bell; su importancia fue inmediatamente reconocida por Robert H. Dicke y su grupo de la Universidad de Princeton. Desde entonces se ha aprendido mucho sobre el espectro de la radiación. Su intensidad se ha estudiado ya en longitudes de onda que se extienden de 30 centímetros hasta medio milímetro, confirmando la conjetura inicial de que su curva espectral se adapta a la de un cuerpo negro a temperatura de tres grados Kelvin.

Uno de los más importantes hechos observados por Penzias y Wilson era la constancia de temperatura de la radiación procedente de distintas direcciones en el espacio. Sus medidas indicaban que la temperatura variaba menos del

10 por ciento en cualquier dirección. Experimentos posteriores establecieron límites todavía más bajos a la desviación respecto a la isotropía. Dos grupos independientes han realizado hace poco mediciones suficientemente sensibles para poner de manifiesto, sin embargo, que la temperatura de la radiación no es exactamente la misma en todas las direcciones. Uno de los conjuntos de experimentos se ha realizado en Princeton por David T. Wilkinson y Brian E. Corey y el otro conjunto en el Laboratorio Lawrence, de la Universidad de California en Berkeley, por un grupo en el que figuran George F. Smoot, Marc V. Gorenstein y yo mismo. Se sabe ahora que la temperatura de la radiación de fondo de tres grados varía en aproximadamente un 10 por ciento de un lado a otro del cielo, estando la parte más caliente en la dirección de la constelación Leo y la más fría en la dirección de Acuario. La temperatura varía suavemente entre esas dos regiones siguiendo una sencilla cosinusoide. Esta forma característica ("el gran coseno del cielo") nos lleva a identificar la velocidad del sistema solar como causa de la anisotropía. Para explicar cómo se ha llegado a esta conclusión y cuál es su importancia, es necesario recordar la teoría de la gran explosión, el origen de la radiación cósmica de fondo y lo que se ha aprendido a partir de la existencia de la anisotropía.

La gran explosión no fue simplemente una explosión de un núcleo de materia en el seno de un vasto espacio que por lo demás estaba vacío. Aunque tal descripción daría cuenta de la ley de Hubble (la observación de que las galaxias distantes se están separando de nosotros a velocidades proporcionales a sus distancias) parece incapaz de explicar la uniformidad con que la materia y la radiación llenan el espacio. El Universo conocido parece estar tan uniformemente poblado, que los astrónomos aceptan el "principio cosmológico": la creencia de que el Universo es esencialmente lo mismo en todas sus partes. Además, la idea de la explosión de un núcleo de materia que se encontrase en algún punto del espacio no ofrece un camino natural para explicar la existencia de la radiación cósmica de fondo. Cualquier radiación emitida en el momento de la explosión habría abandonado la proximidad de la masa original todavía más rápidamente de lo que lo habría hecho la materia, y la radiación no se encontraría ya en el espacio para ser observada.

En la teoría de la gran explosión no hay núcleo primitivo de materia ni centro de la explosión. El espacio está uniformemente ocupado; no hay borde exterior de la distribución de materia. La gran explosión no fue una explosión de materia dentro del espacio, sino una explosión del espacio mismo. Según la teoría de la relatividad general de Eins-



MOVIMIENTO ABSOLUTO DE LA TIERRA a través del espacio. Se ha podido determinar midiendo pequeñas diferencias en la temperatura de la radiación cósmica de fondo de tres grados que llega a la Tierra procedente de diferentes direcciones. La Tierra viaja en su órbita alrededor del Sol a 30 kilómetros por segundo y, como prisionera gravitatoria del Sol, está siendo arrastrada alrededor del centro de la galaxia a 300 kilómetros por segundo. El experimento del nuevo movimiento del éter muestra que el movimiento neto de la Tierra en el espacio es de unos 400 kilómetros por segundo. El vector del movimiento neto de la Tierra está en el mismo plano de su órbita alrededor del Sol y forma un ángulo marcadamente inclinado hacia arriba (hacia el Norte) con el plano de la galaxia. En este diagrama, el vector del movimiento neto de la Tierra se representa mediante una flecha de color centrada en el Sol, ya que ambos cuerpos viajan juntos. Ambos están siendo arrastrados por el propio movimiento "peculiar" de la galaxia a través del espacio (el movimiento peculiar de la galaxia, y no una parte del movimiento cósmico de conjunto). Para explicar el movimiento de la Tierra con respecto a la radiación de tres grados, la galaxia debe estar viajando a unos 60 kilómetros por segundo (más de dos millones de kilómetros por hora) en la dirección que aparece indicada por la flecha negra gruesa.

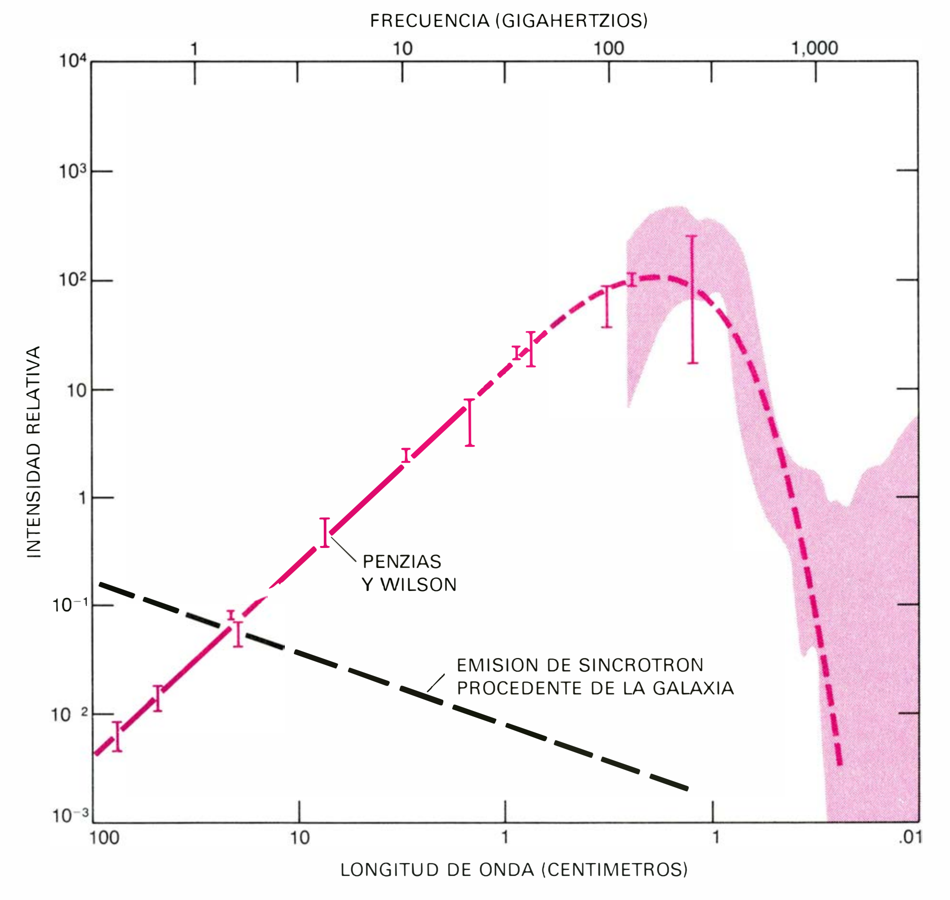
tein, la "cantidad" de espacio entre los objetos no es fija, aun cuando los objetos retengan sus respectivas coordenadas de posición. En los cálculos realizados en la teoría de la gran explosión, se supone generalmente que las galaxias están en reposo aunque la cantidad de espacio entre ellas aumenta. Cualquier movimiento que conduzca a un cambio de las coordenadas de posición de la galaxia en esta teoría se designa como "velocidad peculiar", no porque sea rara, sino porque es peculiar de la galaxia individual y no parte de un movimiento cósmico de conjunto.

La velocidad de expansión del espacio disminuye por la presencia de materia y energía. Si la densidad media del universo es inferior a un valor crítico (alrededor de 10^{-29} gramos por centímetro cúbico), la expansión continuará para siempre. Si la densidad media es menor que el valor crítico, la expansión se retardará hasta detenerse y convertirse en una implosión. La densidad también determina la configuración en gran escala del Universo. Si la densidad es mayor que el valor crítico, el volumen del Universo es finito; en otro caso, el volumen es infinito. Hasta ahora no se ha podido establecer la densidad del Universo con exactitud suficiente para decir inequívocamente si el Universo es finito o infinito. Afortunadamente para la mayoría de los cálculos de la teoría de la gran explosión, esa cuestión no es crucial. Supondremos que la densidad media es igual al valor crítico, lo que tiene la ventaja adicional de implicar que la curvatura media del espacio es nula; por tanto, podemos trabajar con la conocida geometría euclidiana.

La idea de que la distancia entre dos objetos pueda cambiar sin que éstos se muevan parece rara porque es completamente ajena a nuestra experiencia cotidiana. Sin embargo, no es mucho más rara que la de la curvatura del propio espacio. Hay muchos cuentos de hadas y de fantasía científica que describen sucesos en un espacio flexible. Lo que diferencia la teoría general de la relatividad de los meros raptos de imaginación es un conjunto específico de ecuaciones que relacionan la geometría y volumen del espacio con su historia anterior y su contenido de masa y energía.

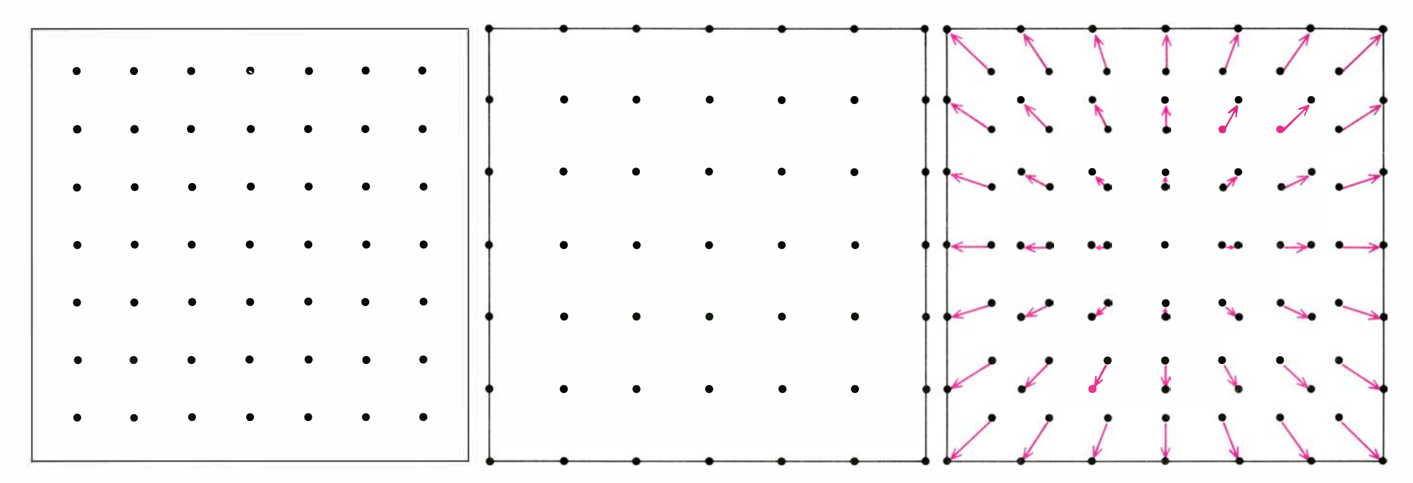
La ley de Hubble encaja de manera natural en la teoría de la gran explosión. La relación es consecuencia de dos hechos: no sólo está el espacio uniformemente ocupado por materia, sino que también se está creando espacio a una velocidad uniforme. Así, cuanto más grande sea la distancia que separa dos galaxias, tanto mayor será la cantidad de espacio creado entre ellas. La observación de Hubble de que todas las galaxias se están alejando de la nuestra no significa que la nuestra esté en el centro del Universo; desde cualquier otra galaxia podría hacerse una observación análoga.

La dilatación uniforme del espacio se aplica solamente a distancias intergalácticas. No se aplica, por ejemplo, en la proximidad de objetos masivos como el Sol, donde la geometría del espacio puede ser completamente diferente. Tampoco se aplica a las distancias entre los átomos en una molécula o de los electrones en un átomo. Tales distancias están determinadas por fuerzas electromagnéticas más que por gravitatorias. Aunque la dilatación del espacio tendiera a separar los constituyentes de átomos y moléculas, sus campos eléctricos internos los acercarían de nuevo. Si este no fuera el caso, los observadores humanos y sus cintas métricas crecerían a la misma velocidad que el Universo, haciendo inobservable la expansión del espacio.

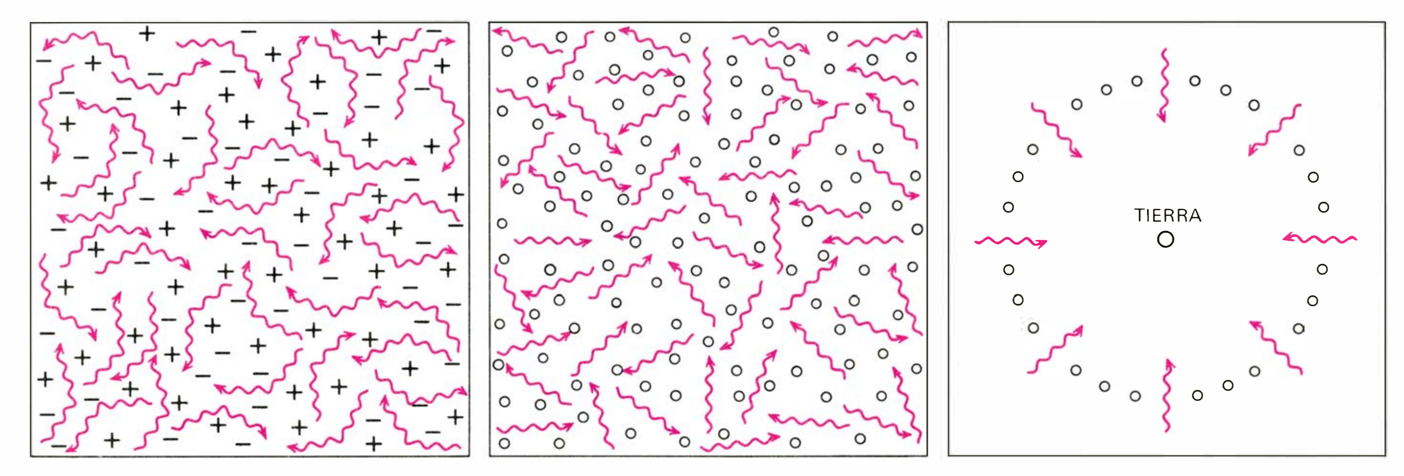


LA INTENSIDAD DE LA RADIACION COSMICA DE FONDO sigue el espectro de energía de un cuerpo negro a la temperatura de tres grados K. La primera medición de esta radiación fue realizada en 1965 por Arno A. Penzias y Robert W. Wilson, trabajando con un receptor de microondas sintonizado a una longitud de onda de 7,35 centímetros (correspondiente a una frecuencia de cuatro gigahertz). La mayor parte de las medidas posteriores se realizaron también en longitudes de ondas fijas, indicadas por las barras verticales. Recientemente, sin embargo, Paul L. Richards y sus colaboradores de la Universidad de California en Berkeley han medido la porción de frecuencia más alta de la curva con una técnica de banda ancha, obteniendo los resultados indicados por la zona coloreada. La línea de trazos representa la radiación de sincrotrón de nuestra galaxia: radiación emitida por los electrones al moverse en torno a las líneas de fuerza magnéticas. A frecuencias por debajo de 10 gigahertz, la anisotropía, o falta de uniformidad direccional, de la emisión de sincrotrón enmascara la anisotropía de la radiación de fondo.

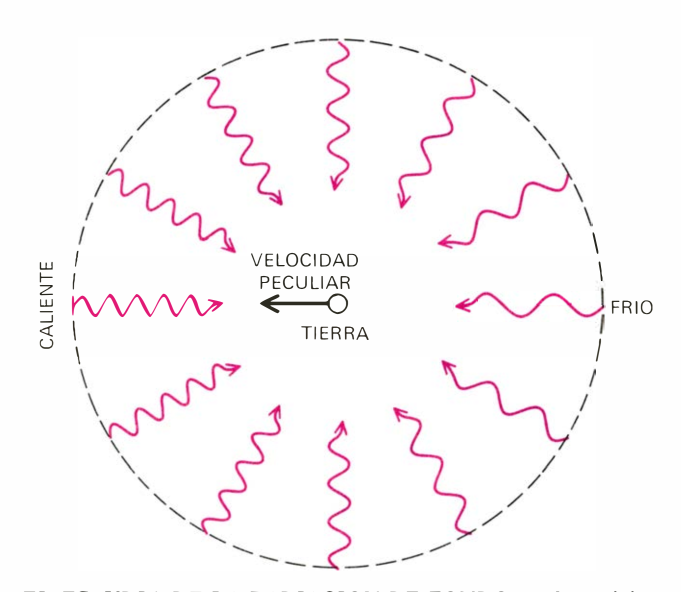
El gran éxito inicial de la teoría de la gran explosión tuvo lugar cuando George Gamov, Ralph A. Alpher y Robert Herman extrapolaron hacia atrás la expansión hasta un tiempo en que la densidad del Universo era más de 10^{30} veces la actual. Postularon que el Universo primitivo debía ser extremadamente caliente y que la combinación de altas temperaturas y densidades iniciaría reacciones termonucleares, convirtiendo el plasma de protones, electrones y neutrones en deuterones y núcleos de helio. Sólo en unos minutos, la expansión del plasma reduciría la temperatura y densidad por debajo del nivel necesario para mantener ulteriores reacciones. La conversión sería incompleta, y exactamente la suficiente para explicar la relación de helio a hidrógeno actualmente existente en el Universo.

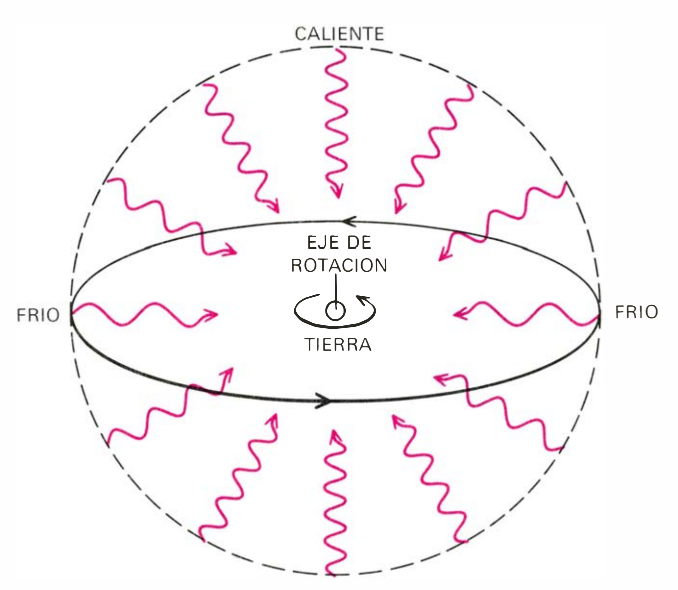


EXPANSION UNIFORME DEL ESPACIO, que explica la "ley" descubierta por Edwin P. Hubble, del Observatorio del Monte Wilson, hace 50 años, cuando observó que las galaxias distantes se están alejando a velocidades proporcionales a sus distancias. Aquí se representa la expansión del espacio por la variación de la distancia entre las galaxias *(puntos)* en las distribuciones de la izquierda y centro. En el diagrama de la derecha se han superpuesto ambas distribuciones. Las flechas que unen los puntos muestran las distancias recorridas por cada galaxia como se vería desde la galaxia central. Aunque el espacio entre las distintas galaxias se expande, el tamaño de cada galaxia no varía.



SEGUN LA TEORIA DE LA GRAN EXPLOSION *(izquierda),* el Universo primitivo estaba lleno de protones *(signo más)* y electrones *(signo menos)* que absorbían y reemitían fotones *(en color).* Tras medio millón de años *(centro),* el Universo se ha expandido y enfriado lo suficiente para que los protones y electrones se combinen formando átomos de hidrógeno *(círculos).* Aquellos fotones *(vueltos a dibujar a la derecha)* que fueron los últimos en difundirse por una capa que rodea la posición en la que se formará la Tierra, constituyen la radiación cósmica de fondo.





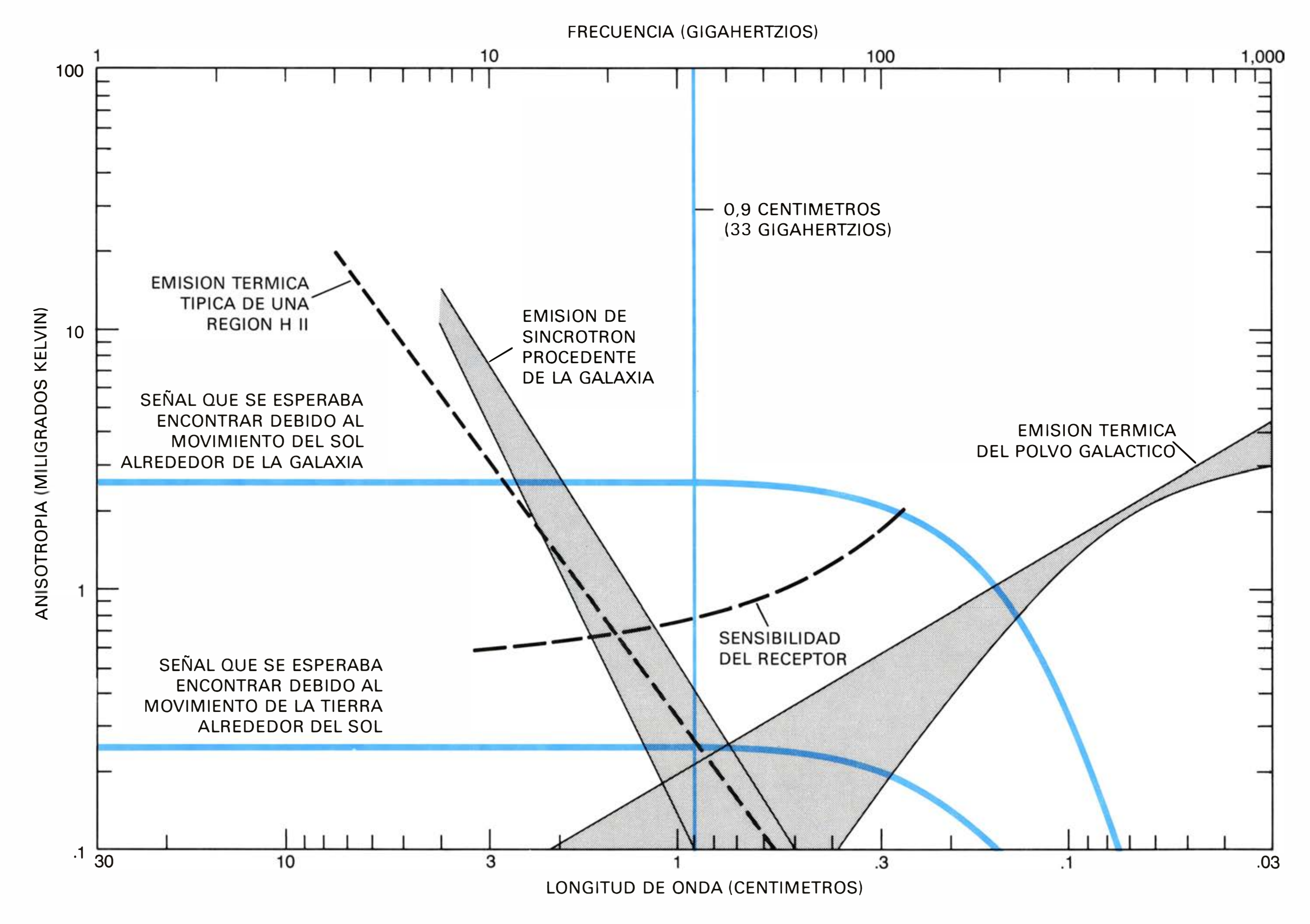
EL ESTUDIO DE LA RADIACION DE FONDO puede suministrar indicios sobre la estructura en gran escala del Universo. Si la Tierra tiene una velocidad peculiar *(izquierda),* la radiación es ligeramente "más azul" (más caliente) en la dirección del movimiento, y "más roja" (más fría) en la opuesta. Si la capa de materia que por última vez difundió la radiación estaba en rotación con respecto a nuestro sistema de referencia local e inercial, los fotones emitidos en el ecuador de la capa se ven retardados por su velocidad adicional (según la teoría general de la relatividad). Las dos posibilidades se pueden distinguir por diferencias registradas en la distribución de temperaturas observadas. La temperatura de la radiación variaría en el primer caso como en el coseno del ángulo en el cielo, y en el segundo caso como el cuadrado del coseno.

Otra consecuencia del modelo de Gamow-Alpher-Herman, que en aquel momento pasó inadvertida, era que el plasma caliente emitiría y absorbería radiación electromagnética de igual manera que el plasma caliente en la superficie del Sol emite luz. La radiación sería difundida y vuelta a difundir por los electrones libres hasta aproximadamente medio millón de años tras la gran explosión. En esa época, la densidad y temperatura de la materia habrían descendido hasta el punto en que sus iones constitutivos (en su mayor parte protones y electrones) se unirían para formar átomos eléctricamente neutros. Este período (que en realidad duró varios miles de años) se llama generalmente "el momento del desacoplo", porque desde entonces en adelante hay poca interacción entre radiación y materia. El Universo, que antes era opaco, se hace de golpe transparente, permitiendo que la radiación electromagnética viaje por el espacio sin difundirse y conservando la imagen que el plasma ofrecía cuando se difundieron los fotones por última vez.

Esta es la radiación que ahora observamos como radiación cósmica de fondo. La radiación que hoy nos llega a nosotros fue difundida, por última vez, por una envoltura de plasma que rodeaba completamente nuestra actual posición en el espacio. Si parte de la materia existente en ese plasma se ha reagrupado en una galaxia muy lejana a la nuestra, podemos imaginar que si en ella hay seres inteligentes, éstos están observando ahora la radiación que fue difundida en nuestra región del espacio hace quince mil millones de años.

Emitida originariamente como radiación visible e infrarroja con una longitud de onda de 0,7 micrómetros para el máximo de energía, la radiación cósmica de fondo ha sufrido un corrimiento hacia el rojo que multiplica las longitudes de onda por 1500, de modo que ahora observamos la longitud de onda correspondiente al máximo de energía en aproximadamente un milímetro. El corrimiento hacia el rojo se debe a la velocidad, enormemente alta, de la envoltura radiante en expansión, o, más propiamente, la alta velocidad a la cual está aumentando el espacio entre nosotros y la envoltura de radiación. La radiación en sí no ha alterado su longitud de onda. Lo que ocurre es que la estamos observando desde un sistema de referencia que se está "moviendo" al 99,9 por ciento de la velocidad de la luz respecto a la materia que la emitió hace quince mil millones de años.

Una notable característica del espectro de un cuerpo negro es que, cuando se observa desde un sistema de referencia móvil respecto al emisor, retiene su forma característica de espectro de cuerpo negro, alterándose sólo la temperatura. En un sistema de referencia móvil con el plasma, la temperatura característica de la radiación es de unos 4500 grados Kelvin; en nuestro sistema de referencia,



LA DETERMINACION DE LA ANISOTROPIA en la radiación de fondo se complica por la presencia de emisión en microondas procedente de varias fuentes que son en sí mismas anisótropas. Las regiones H II, por ejemplo, son concentraciones de gas y polvo calentado por estrellas jóvenes. Para el experimento del nuevo movimiento del éter fue preciso elegir una frecuencia a la que la supuesta anisotropía de la radiación de fondo predominase. Se consideró óptima una frecuencia de 33 gigahertz. Las dos curvas que representan las señales que se esperaba encontrar a causa del movimiento de la Tierra alrededor del Sol (30 kilómetros por segundo) y del movimiento del Sol alrededor del centro de la galaxia (300 kilómetros por segundo) se han calculado en la hipótesis de que la velocidad sea nula para la galaxia.

es de tres grados. A medida que pase el tiempo, continuaremos interceptando radiación cósmica de fondo, pero la señal que entonces observaremos procederá de regiones todavía más distantes en el espacio. Como esas regiones más distantes se están alejando a velocidades todavía mayores, en nuestro sistema de referencia apreciaremos que la radiación tiene una temperatura inferior a tres grados. En otros quince mil millones de años, la radiación que alcance nuestra actual posición en el espacio corresponderá a una temperatura de 1,5 grados. Será también radiación emitida en el momento del desacoplo, pero procedente de una región mucho más distante en el espacio que la radiación que observamos hoy.

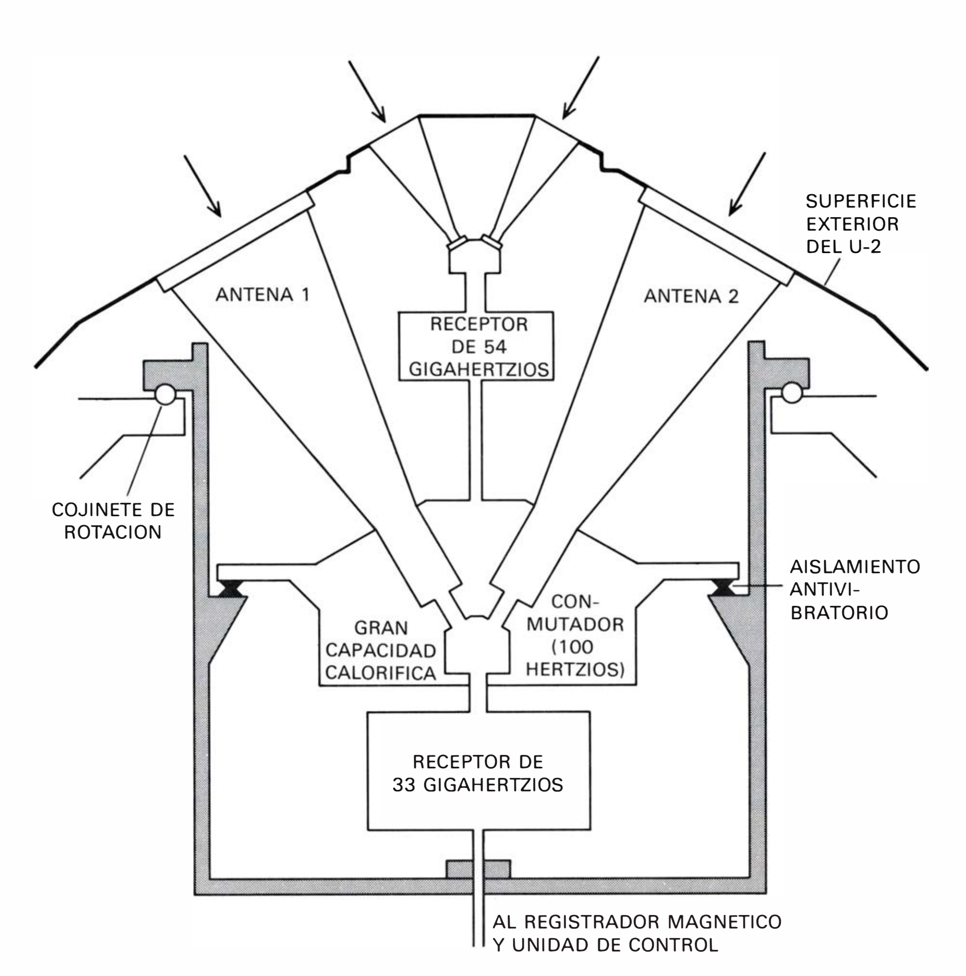
Cuando se observa la radiación cósmica de fondo, se está indirectamente estudiando la estructura de la capa de materia que la difundió medio millón de años después de la gran explosión. Si el Universo fuera totalmente homógeneo e isótropo, la señal de la radiación cósmica carecería de características especiales. Claramente, el actual Universo difiere de la homogeneidad, conteniendo como lo hace planetas, estrellas, galaxias y grupos de galaxias. Si el agrupamiento en gran escala hubiera comenzado antes del momento del desacoplo, la radiación de fondo debería exhibir regiones brillantes y obscuras correspondientes a los agrupamientos de materia. Si se pudieran observar tales características, se obtendría una visión fascinante de la primitiva evolución del Universo. Por otra parte, la ausencia de tales caracteres indicaría que el agrupamiento en gran escala, tal como es necesario para explicar los grupos de galaxias, no se había producido todavía en el momento del desacoplo.

La radiación de fondo proporciona también una oportunidad para someter a prueba algunas de las teorías más especulativas del Universo. Por ejemplo, el Universo puede estar en rotación, posibilidad permitida por la teoría general de la relatividad. S. W. Hawking, de la Universidad de Cambridge, fue el primero en llamar la atención sobre el hecho de que la rotación se manifestaría claramente como una desviación especial respecto de la isotropía en la radiación cósmica de fondo. Si la capa productora de la última difusión estuviera en rotación con respecto a nuestro sistema local e inercial de referencia, el plasma del ecuador de tal capa tendría una velocidad transversal que no aparecería en el plasma de sus polos. Según el efecto de dilatación del tiempo de la teoría de la relatividad especial, los relojes y otros osciladores a lo largo del ecuador de la capa de plasma se retrasarían, con el resultado de que la luz emitida por la región ecuatorial sufriría un pequeño corrimiento hacia el rojo, además del debido al distanciamiento de la capa de plasma. El corrimiento adicional daría por resultado una temperatura ligeramente más baja para la radiación procedente de la región ecuatorial.

Aunque según la teoría general de la relatividad, un Universo en rotación sería detectable, no lo sería de acuerdo con un principio formulado por Ernst Mach. Mach postuló que la existencia misma de sistemas locales e inerciales de referencia dependía de la materia distante en el Universo. Así, un sistema de referencia local e inercial estaría inextricablemente ligado a la materia distante, y estaría en rotación si el Universo en conjunto lo estuviera. Si el análisis de la radiación de fondo demostrase que el Universo está en rotación, el principio de Mach resultaría ser falso.

Si durante los momentos iniciales de la gran explosión se hubieran generado olas gravitatorias de longitud de onda muy grande, también ellas darían lugar a unos caracteres distintivos en la radiación cósmica de fondo. (Como el momento de desacoplo para olas gravitatorias se produce sólo una fracción de segundo después de la gran explosión, su detección directa suministraría una ojeada a la historia del Universo en momento todavía anterior al que proporciona la radiación cósmica de fondo.) Se podrían descubrir también anisotropías que revelasen que el Universo no se ha expansionado uniformemente, de estricto acuerdo con la ley de Hubble. Tales fenómenos tenderían a formar diferentes configuraciones en el cielo, haciendo posible distinguirlos entre sí. Tal vez el rasgo más característico a buscar, sin embargo, sea la anisotropía causada por el movimiento del sistema solar respecto a la corteza de plasma que emitió la radiación.

En cada región del espacio sólo puede



INSTRUMENTO PARA MEDIR LA ANISOTROPIA de la radiación cósmica de fondo construido por el autor y sus colegas. Aquí está representado esquemáticamente en corte vertical. Las dos grandes antenas de bocina se proyectaron para recoger la radiación cósmica de fondo en un estrecho cono a una frecuencia de 33 gigahertz. Las dos bocinas más pequeñas y su receptor asociado vigilan la emisión del oxígeno atmosférico a 54 gigahertz. El aparato se ideó no para medir la temperatura absoluta de la radiación cósmica de fondo, sino más bien la diferencia entre las temperaturas de las señales recogidas por las dos bocinas grandes cuando se conectan alternativamente con un receptor común 100 veces por segundo. Para compensar asimetrías en el proyecto y la construcción, el aparato se hace girar 180 grados cada 64 segundos en la toma de datos.

haber un sistema inercial de referencia en el que la radiación de fondo sea completamente isotrópica. En cualquier otro sistema, el movimiento del observador se manifestará como una variación de la temperatura de la radiación proporcional a la velocidad del observador y al coseno del ángulo entre su dirección de movimiento y la dirección de la observación. P. J. E. Peebles, uno de los físicos del grupo de Dicke que identificó correctamente la radiación, acuñó la frase "nuevo movimiento del éter" para describir el movimiento esperado. Aunque no es en realidad movimiento con respecto a un sistema de referencia fijo en el espacio, es movimiento con relación al más natural sistema de referencia en cosmología: el sistema de coordenadas en expansión con respecto al cual las galaxias estén casi en reposo.

Fue precisamente el darnos cuenta de la posibilidad de detectar el nuevo movimiento del éter lo que nos inspiró a mis colegas y a mí para idear un experimento que representara una considerable mejora sobre mediciones anteriores. Esperábamos descubrir que el movimiento de la Tierra era principalmente debido al movimiento del sistema solar alrededor del centro de nuestra galaxia a unos 300 kilómetros por segundo, modificado por un pequeño factor para tener en cuenta el movimiento de la galaxia hacia la de Andrómeda. (El movimiento relativo de nuestra galaxia y la de Andrómeda se había medido previamente mediante la variación Doppler de las líneas de emisión espectral, resultando de 80 kilómetros por segundo.) Sólo una pequeña parte del movimiento del éter que se esperaba encontrar se debería al movimiento de la Tierra alrededor del Sol, a 30 kilómetros por segundo.

¿Por qué estábamos tan excitados cuando íbamos a medir una cantidad tan bien conocida? Nuestro principal interés estribaba en otros posibles efectos: la rotación del universo, indicios prematuros de la formación de grupos de galaxias, olas gravitatorias y una expansión de Hubble anisótropa. Al comenzar un experimento difícil, sin embargo, es tranquilizador saber que uno va a obtener un valor no nulo de alguna clase. Aun cuando los otros fenómenos sean interesantes, e incluso un valor nulo como resultado de ellos sería importante, es desalentador hacer mediciones precisas del cero.

Cuando iniciamos el experimento, se sabía que la radiación cósmica de fondo era isótropa dentro de unos pocos miligrados, es decir, con un error inferior a una parte en 500, debido en gran parte a las cuidadosas mediciones de Wilkinson y Robert B. Partridge, de Princeton, y Edward K. Conklin, de la Universidad Stanford. Otro experimentador de Princeton, Paul Henry, había detectado una pequeña desviación de la isotropía, pero sus datos no se ajustaban a una curva sencilla y la dirección de la temperatura máxima no se había determinado con exactitud.

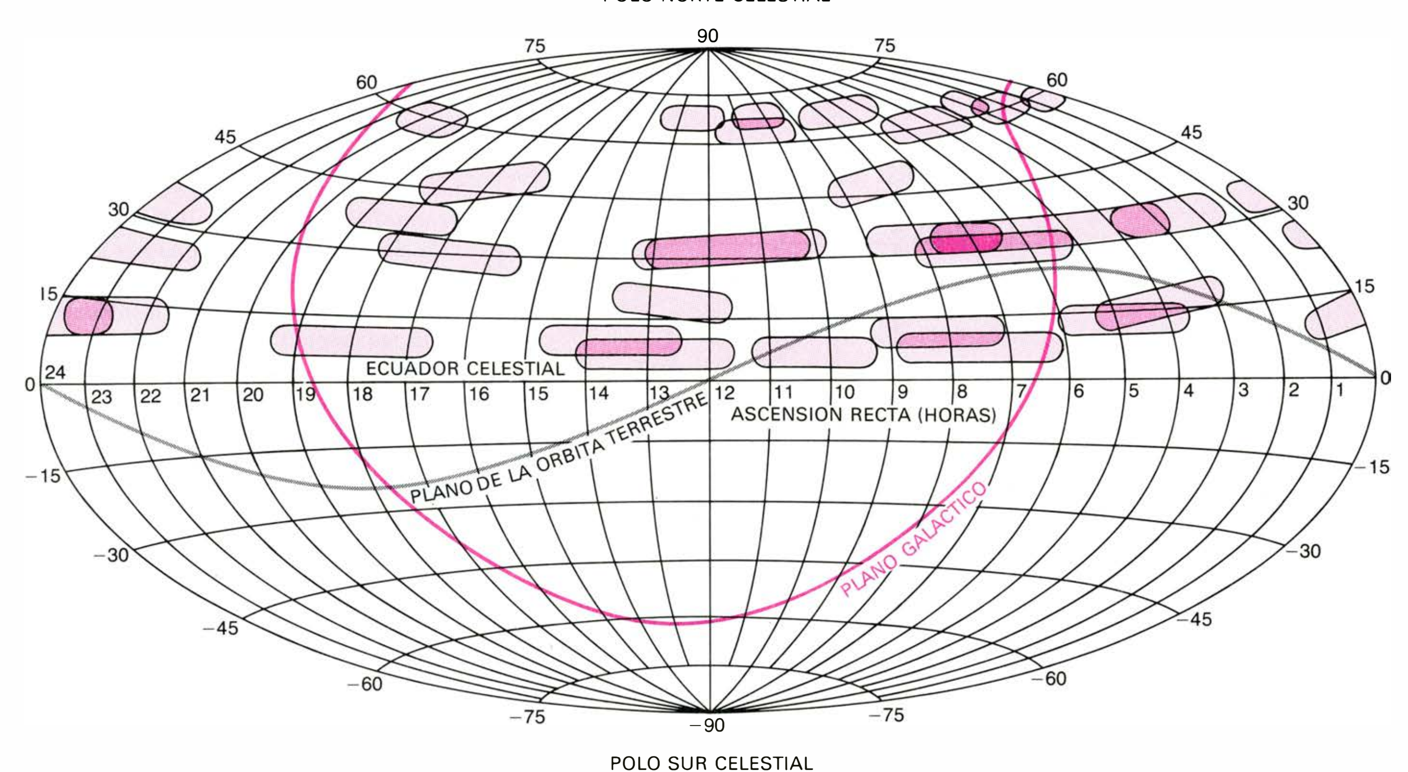
Para nuestras mediciones, proyectabamos usar un instrumento que en general tenía las mismas características que los usados en estudios previos: un radiómetro Dicke. Con este instrumento no se mide la temperatura absoluta de la radiación cósmica, sino diferencias de temperatura entre una dirección en el cielo y otra. Aunque se podría tratar de medir tales diferencias comparando la salida de dos receptores que apuntan en diferentes direcciones, el ruido térmico en los dos receptores y ciertas variaciones incontrolables en su ganancia ("ruido de centelleo") enmascararían las pequeñas diferencias que se esperaban. En el modelo de Dicke se evita el problema conmutando alternativamente el mismo receptor a una y otra de dos antenas en forma de bocina apuntadas en direcciones diferentes. Si el experimento se lleva a cabo en la superficie de la Tierra, el operador trata de anular la intensa emisión en microondas procedentes del oxígeno en la atmósfera apuntando las dos bocinas al mismo ángulo zenital, de manera que ambas "vean" la misma cantidad de oxígeno.

Para eliminar pequeñas diferencias en el poder receptor de las dos bocinas o una posible asimetría en el conmutador de microondas que conecta las bocinas al receptor, se hace girar todo el conjunto del aparato de manera que las posiciones de las dos bocinas se intercambien una vez por minuto. Con esas precauciones, cualquier asimetría presente en la radiación cósmica de fondo debería aparecer como fluctuaciones en la salida del receptor coincidentes con el ritmo de cambio de bocina.

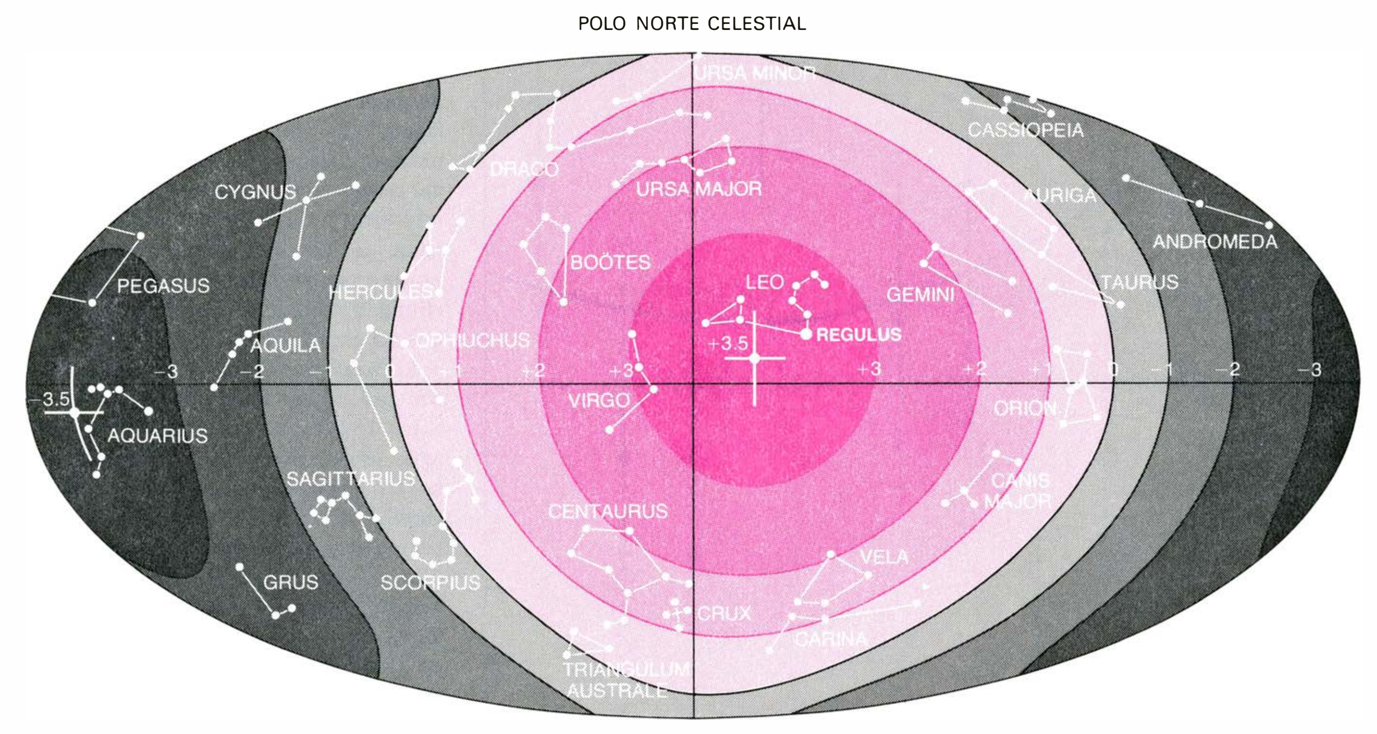
A fin de lograr una mejora substancial

PROBLEMA	REMEDIO
Emisión de sincrotrón procedente de la galaxia	Hacer mediciones por encima de 10 gigahertz
Emisión procedente de polvo galáctico	Hacer mediciones por debajo de 100 gigahertz
Emisión del vapor de agua atmosférico	Tomar datos a altitud superior a 15 kilómetros (con el U-2)
Emisión debida al oxígeno atmosférico	Usar dos antenas de bocina a gran altura y vigilar la emisión del oxígeno en 54 gigahertz
Emisión del Sol	Volar de noche
Emisión de la Tierra	Usar antena doble protegida y con campo estrecho
Emisión de las antenas de bocina	Hacer simétrica la emisión mediante cuidadoso control de la temperatura
Ruido térmico en el receptor	Integrar la señal durante 20 minutos
Ruido de centelleo en el receptor (= 1/frecuencia)	Conmutar entre las dos antenas 100 veces por segundo
Asimetría del avión	Invertir dirección del vuelo cada 20 minutos
Asimetría en el aparato	Girar aparato de medida cada 64 segundos
Asimetría introducida por campo magnético terrestre	Blindar cuidadosamente conmutadores de microondas
Emisión de radio del U-2	Colocar protección metálica alrededor de las partes sensibles y reducir al mínimo las comunicaciones con el U-2
Distorsión geométrica de la atmósfera debida a falta de esfericidad de la Tierra	Determinar el zenit a partir de la señal del oxígeno y no del horizonte terrestre o instrumentos de vuelo

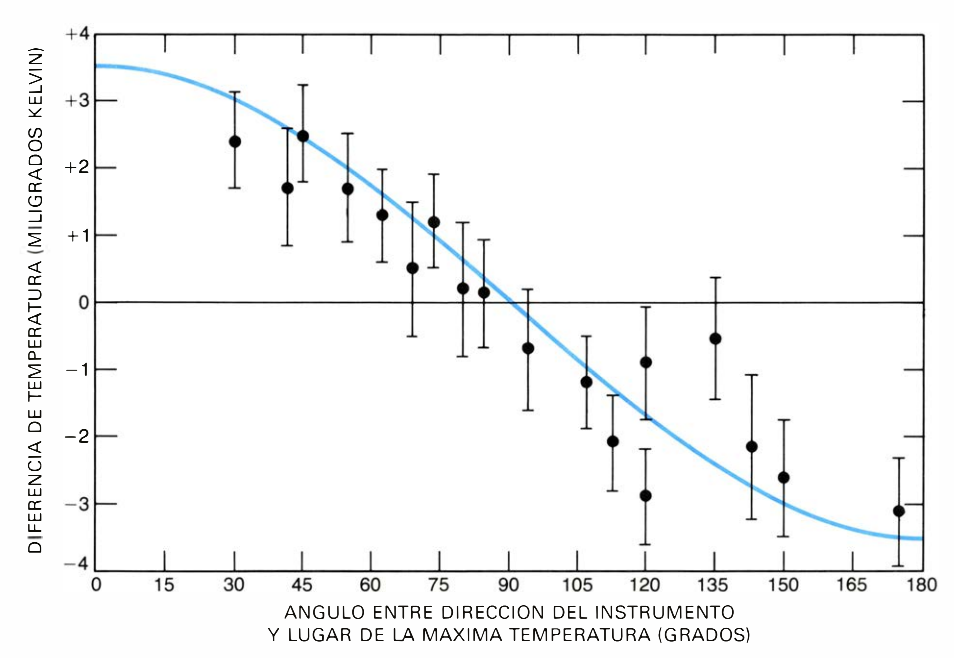
LISTA DE LOS PRINCIPALES PROBLEMAS asociados a la medición de la anisotropía de la radiación de fondo, junto con los remedios adoptados por los investigadores. Estos y sus colaboradores pasaron tres años planeando y construyendo el equipo necesario antes del vuelo de prueba.



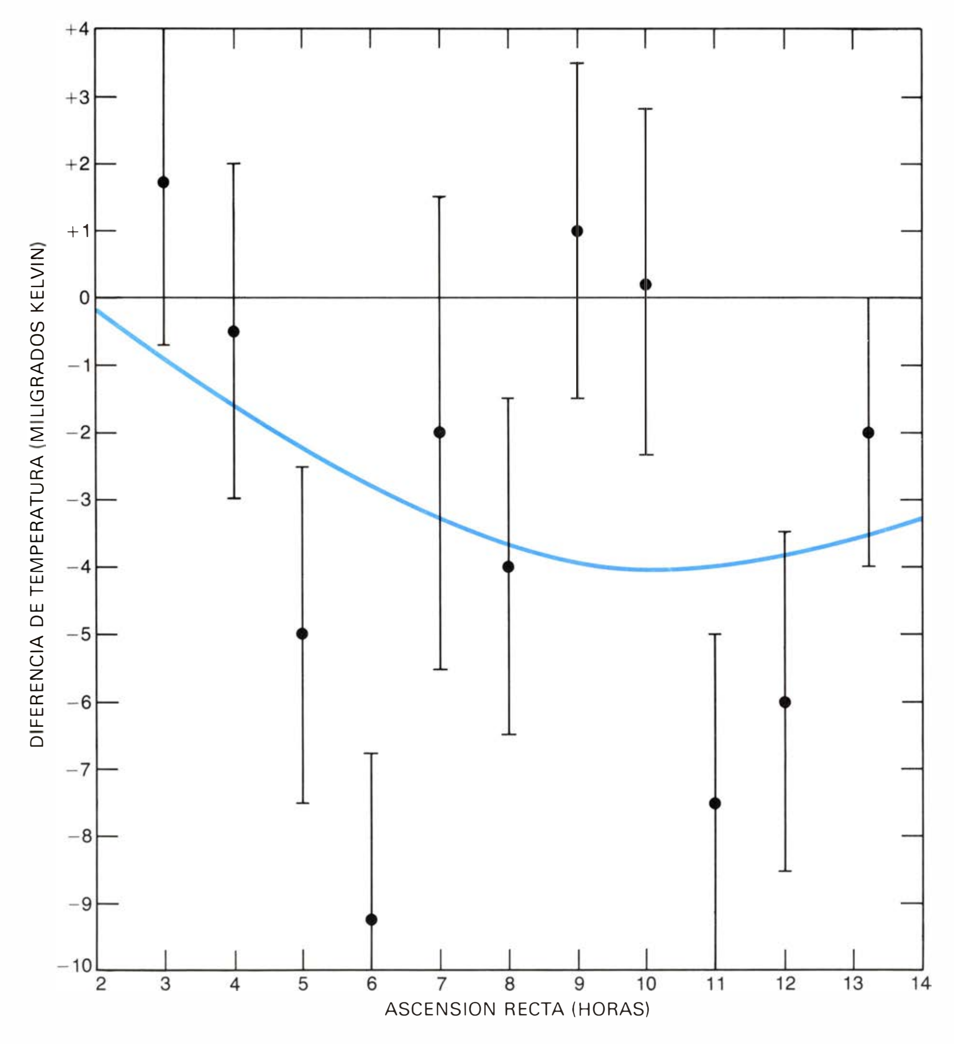
REGIONES DEL CIELO EXPLORADAS por el experimento del nuevo movimiento del éter, representadas sobre una proyección de la esfera celeste que conserva las áreas. Todos los vuelos se hicieron de noche (para evitar la emisión en microondas del Sol) y se extendieron a lo largo de un período de aproximadamente un año (para explorar la mayor porción posible del cielo que puede observarse desde una estación en el Norte de California). Los vuelos en el Hemisferio Sur, que se proponen para el año en curso, ayudarán a comprobar el aspecto de la anisotropía de temperatura detectada. Los vuelos típicos duraban cuatro horas y exploraban cuatro regiones diferentes del cielo. Cada una de las antenas de bocina era sensible a una región del cielo de unos siete grados de diámetro aproximadamente. La longitud de cada zona explorada estaba determinada principalmente por la rotación de la Tierra más que por la velocidad (unos 750 kilómetros por hora, aproximadamente) del avión.



ANISOTROPIA de la radiación de fondo, representada sobre la esfera celeste en intervalos de un miligrado Kelvin. El punto "más caliente", indicando la dirección del máximo movimiento relativo de la Tierra hacia la radiación de fondo, se halla en la constelación Leo a la ascensión recta de 11 horas (±0,5 horas) y seis grados de latitud (±10 grados). El punto "más frío" o dirección en la cual la radiación está más "enrojecida" por el movimiento relativo de la Tierra alejándose de los fotones incidentes, se encuentra a 180 grados de distancia angular. Si la diferencia de temperatura entre los puntos más calientes y más frío se representa en función de la distancia, resulta una cosinusoide.



LA COSINUSOIDE proporciona el mejor ajuste para los datos (promediados en 18 puntos) tomados por el autor y sus colegas en el experimento del nuevo movimiento del éter. El eje horizontal representa el ángulo formado por la línea que une las dos antenas de bocina y la dirección del máximo de temperatura en Leo. La cosinusoide es la distribución de temperatura que podría esperarse si la velocidad peculiar del sistema solar hacia Leo fuese de 400 kilómetros por segundo.



PRIMERA DESVIACION IMPORTANTE respecto a la isotropía en la radiación cósmica de fondo; fue detectada por Paul Henry, de la Universidad de Princeton, con un instrumento que se elevó en un globo. La anisotropía de la radiación se manifiesta en la preponderancia de puntos que caen por debajo de la línea de cero. La dispersión de los puntos, sin embargo, imposibilitó descubrir la distribución de la anisotropía o determinar la dirección del máximo de temperatura.

de sensibilidad sobre experimentos anteriores, teníamos que saber exactamente lo que había limitado la sensibilidad de aquéllos y prever de la mejor manera posible las dificultades que una nueva concepción experimental introduciría. Los resultados de experimentadores anteriores habían sido reducidos en su sensibilidad primordialmente por la llamada "radiación de sincrotrón", emitida por electrones acelerados en el campo magnético de nuestra galaxia. Aunque la intensidad de la radiación de sincrotrón sigue aproximadamente los caracteres visibles de la Vía Láctea, su aspecto preciso en la región de microondas del espectro no se conoce. Lo mejor que se puede hacer es sustraer de la anisotropía total observada una estima de la anisotropía de la radiación de sincrotrón y esperar que la diferencia represente la anisotropía de la radiación cósmica de fondo y no el resultado de una estima errónea.

Hay un camino directo para reducir la interferencia introducida por la radiación de sincrotrón, que consiste en trasladarse a longitudes de onda inferiores a tres centímetros. Por ejemplo, entre tres centímetros y un centímetro la radiación de sincrotrón se reduce aproximadamente a una tercera parte. De igual importancia es que la radiación cósmica de fondo, que sigue la curva de emisión del cuerpo negro, en el mismo intervalo de longitud de onda se hace aproximadamente 10 veces más intensa. El obstáculo que se encuentra al operar a longitudes de onda más cortas es la emisión atmosférica creciente: el vapor de agua y el oxígeno hacen imposibles las observaciones desde Tierra para longitudes de onda inferiores a dos centímetros. El vapor de agua es particularmente perturbador, porque puede darse en bolsas que no se eliminan apuntando un par de antenas al mismo ángulo zenital.

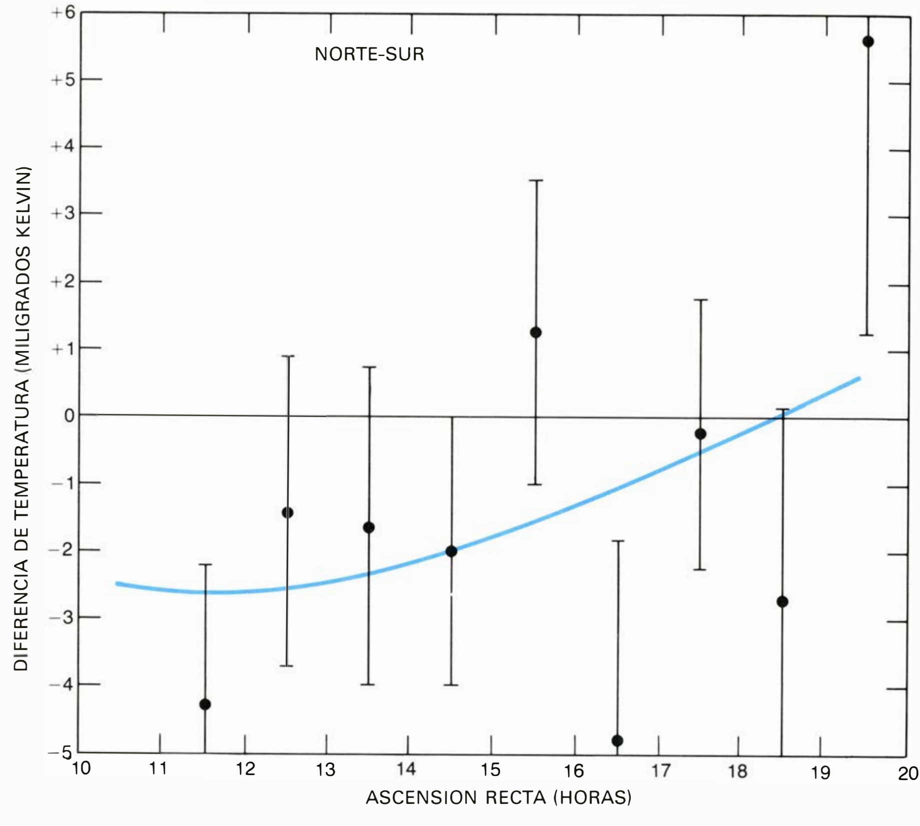
La solución obvia es realizar el experimento a una altura bastante superior a los 15.000 metros, donde el vapor de agua se ha congelado casi completamente. La cima de una montaña no es suficiente. Es necesario usar bien un globo o un avión, o un vehículo espacial. Aunque sabíamos que el experimento en vehículo espacial era potencialmente el más sensible, un experimento en avión o en globo es mucho más barato y ciertamente debe realizarse antes. Al discutir estos problemas con Hans M. Mark, entonces director del Centro de Investigación Ames de la Administración Nacional de Aeronáutica y del Espacio (NASA), y Luis W. Alvarez, del Laboratorio Lawrence de Berkeley, decidimos que el avión U-2, que estaba siendo utilizado por la

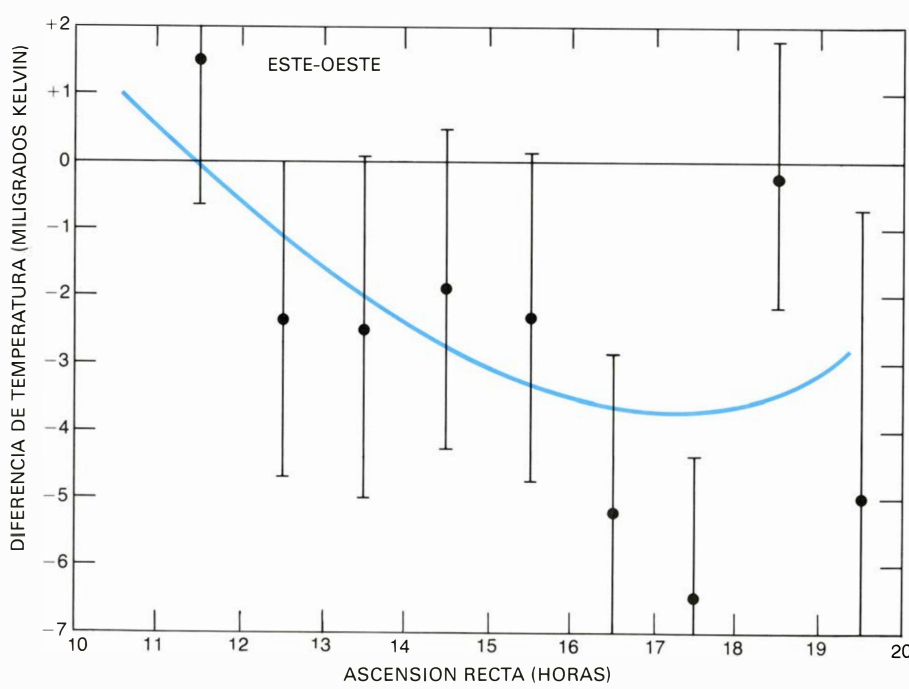
NASA para el estudio de los recursos terrestres, sería una plataforma ideal para nuestro experimento. Hacia la misma época (a mediados de 1973) Corey y Wilkinson, en Princeton, decidieron usar la barquilla de un globo como plataforma para sus experimentos de anisotropía. No voy a describir su experimento, concentrándome en lugar de ello en los problemas que había que resolver para nuestro proyecto U-2.

Para un experimento realizado en el aire, el tiempo diponible está severamente limitado, lo que significa que nuestro receptor tenía que tener la máxima sensibilidad posible. (En un experimento realizado desde el suelo, los datos se pueden promediar para muchas observaciones y obtener resultados válidos aun con un receptor de bajo ruido.) Desgraciadamente, los receptores de microondas se hacen progresivamente menos sensibles a longitudes de onda inferiores a los tres centímetros. Las limitaciones que imponía la tecnología de receptores, la necesidad de evitar la perturbadora interferencia de la radiación de sincrotrón a longitudes de onda mucho mayores de 1,5 centímetros y de fuertes líneas atmosféricas de emisión a longitudes de onda de algunos milímetros nos condujo finalmente a elegir una longitud de onda de 0,9 centímetros (una frecuencia de 33 gigahertz o 33.000 millones de ciclos por segundo) como la óptima para el experimento. A esa longitud de onda creíamos que nuestro aparato sería suficientemente sensible para detectar una anisotropía de menos de una milésima de grado, lo que sería más que adecuado para determinar la velocidad a la cual el sistema solar va siendo barrido alrededor de disco galáctico.

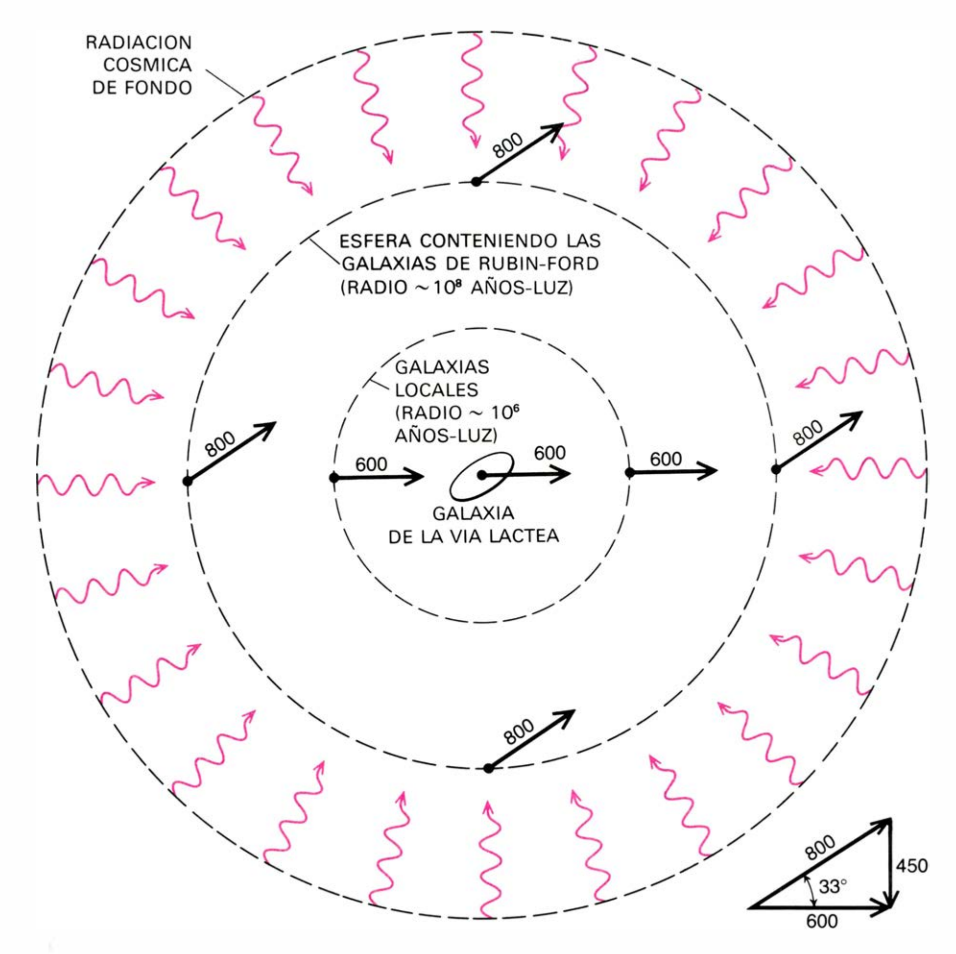
En un experimento en avión constituye un grave problema la inestabilidad de la plataforma, que hace difícil asegurar que las dos antenas de bocina estén apuntando al mismo ángulo zenital y, consiguientemente, observando la misma cantidad de oxígeno. Resolvimos el problema observando el ángulo zenital con un segundo radiómetro sintonizado a la longitud de onda de 0,55 centímetros (una frecuencia de 54 gigahertz), longitud de onda particularmente sensible a la emisión del oxígeno atmosférico. Con este dispositivo estaríamos en condiciones de detectar cualquier asimetría de la señal del oxígeno, tanto si se debía a inclinación del avión como a inclinación de la atmósfera. Como la Tierra no es una esfera, sino un cuasi-elipsoide, la atmósfera está en realidad frecuentemente "inclinada" respecto al suelo; la atmósfera también está inclinada en los frentes meteorológicos. Si la inclinación fuese grande, se pediría al piloto que ladease el avión para compensar. (La maniobra no resultó necesaria.)

El tamaño de nuestro aparato estaba estrictamente limitado por el espacio disponible dentro de la bodega posterior del U-2, lo que hizo el proyector de las antenas particularmente difícil. Como la Tierra es una intensa fuente emisora de microondas, teníamos que dar con una manera de resguardar nuestras antenas. En los experimentos desde el suelo, el problema se había resuelto mediante

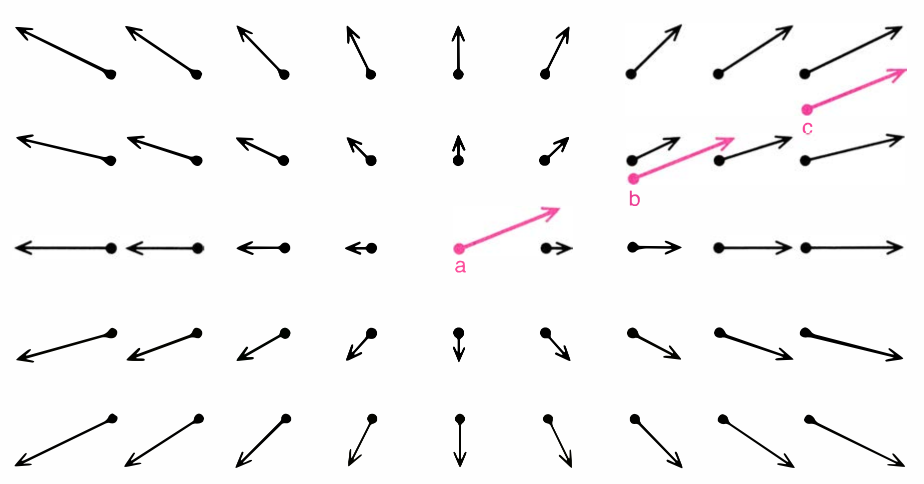




EL SEGUNDO EXPERIMENTO DE PRINCETON, bajo la dirección de David T. Wilkinson y Brian E. Corey con un instrumento montado en un globo y operando en la frecuencia de 19 gigahertz, confirma las mediciones del U-2 de Berkeley. Como los investigadores de Princeton representaron sus datos en forma diferente (mediante una proyección Norte-Sur y otra Este-Oeste), la semejanza entre sus resultados y los de Berkeley no aparece a primera vista. Está claro, sin embargo, que los datos de Princeton definen cosinusoides. El grupo de investigadores de la Universidad de Princeton llegó a la conclusión de que la Tierra se está moviendo a 300 (±70) kilómetros por segundo hacia la ascensión recta 12 (±2) horas y la latitud -10(±20) grados en la esfera celeste.



LA FANTASTICA DISTRIBUCION DE VELOCIDADES que aparece cuando la velocidad peculiar de nuestra galaxia, evidentemente compartida por todos los miembros del grupo local de galaxias, se representa respecto a una muestra de galaxias a 10^8 años-luz cuyas velocidades fueron determinadas espectrográficamente por Vera C. Rubin y W. Kent Ford, Jr., del Departamento de Magnetismo Terrestre de la Institución Carnegie de Washington. De sus resultados se deduce que nuestra galaxia se está moviendo a 450 kilómetros por segundo respecto a las que constituyen la muestra de referencia. El diagrama revela cómo la velocidad de Rubin-Ford se puede reconciliar con la peculiar de 600 kilómetros por segundo determinada para nuestra galaxia por la anisotropía de la radiación cósmica de fondo. La esfera de galaxias de Rubin-Ford necesitaría una velocidad peculiar de 800 kilómetros por segundo desviada unos 33 grados de la dirección en que se está moviendo nuestra galaxia. El diagrama de la derecha muestra cómo nuestra galaxia se estaría aproximando a las de Rubin-Ford a 450 kilómetros por segundo. En vista de la incertidumbre por imperfección en las mediciones, las velocidades se han redondeado a 50 kilómetros por segundo.



CON EL TIEMPO, puede esperarse la conversión de velocidad peculiar en la velocidad de Hubble, que es la velocidad cósmica de expansión. La flecha *a* representa el vector velocidad peculiar actual de nuestra galaxia, y aparece sumergido en un espacio que se está expandiendo uniformemente. A medida que nuestra galaxia se mueve hacia el exterior, alcanzará otras galaxias *(b)*, hasta que llegue a una región *(c)* en que su velocidad sea igual a la de los objetos circundantes. A partir de entonces nuestra galaxia no tendrá velocidad peculiar; su movimiento respecto a la materia próxima tenderá a cero. Un argumento similar muestra que, en el pasado, la velocidad peculiar de nuestra galaxia debe haber sido mayor de lo que es hoy. Esta línea de argumentación queda invalidada si la velocidad peculiar aparece como consecuencia de un efecto local dado.

grandes reflectores metálicos que interceptaban y reflejaban el resplandor terrestre. Nuestra solución fue una bocina de antena especial, proyectada de modo que tuviera una capacidad de recepción extremadamente pequeña para direcciones que formasen un ángulo mayor de 60 grados con el eje de la bocina. El pequeño espacio disponible en el U-2 requería también que el aparato fuese enteramente automático, ya que no había sitio en el avión para un científico pasajero. Otro problema no sencillo era que el U-2 se construyó para llevar cámaras que apuntasen hacia abajo y nosotros queríamos apuntar hacia arriba. No se puede cortar un agujero en la parte superior del fuselaje de un avión de lámina pretensada como el U-2 sin hacer muchos cálculos, pero se logró llevar a cabo la modificación con la ayuda del personal de NASA en el Centro de Investigación Ames y los ingenieros de la Lockheed Aircraft Corporation, que era responsable del mantenimiento del avión.

Eso fue sólo una muestra de los problemas que hubimos de resolver en el experimento. Parte del interés de un experimento nuevo es tratar de anticipar nuevos problemas y de resolverlos. Mucho del mérito debe atribuirse a mis colaboradores Smoot y Gorenstein, cuya principal responsabilidad era transformar un proyecto teórico en un experimento con éxito.

Tras tres años de planeamiento, construcción y pruebas, montamos el aparato en el U-2 en julio de 1976. Hicimos varias modificaciones después de una serie de vuelos de prueba y continuamos haciéndolas durante el período de toma de datos, que comenzó en diciembre de ese año. Todos los vuelos de toma de datos se realizaron por la noche porque nuestras antenas especiales de bocina detectaban señales en microondas procedentes del Sol. No había tampoco manera práctica de proteger nuestro aparato del calentamiento solar irregular. La emisión en microondas de la Luna, cuando se encontraba en el ángulo correcto para ello, proveyó una cómoda manera de calibrar la ganancia del receptor en vuelo.

Los datos recogidos desde el primer vuelo mostraron una inequívoca desviación respecto a la isotropía en la radiación cósmica de fondo. Para obtener un cuadro claro de la anisotropía, sin embargo, teníamos que tener vuelos que se extendiesen a lo largo de todo el año, de modo que las antenas pudiesen explorar una fracción lo mayor posible de la esfera celestial visible desde el Norte de California. Hacia el final del tercer año,

los datos de diez vuelos configuraban claramente la curva cosinusoidal que se podría esperar si el sistema solar se estuviera moviendo con una alta velocidad cosmológica. El radiómetro instalado en la barquilla de un globo por Corey y Wilkinson, operando a 19 gigahertz, detectó una anisotropía análoga a 1,6 centímetros. Tanto en el experimento de Berkeley como en el de Princeton, la magnitud de la anisotropía concordaba con la previamente hallada por Henry.

Nuestros datos indican que la temperatura de la radiación cósmica de fondo alcanza un máximo de 0,0035 grados (3,5 miligrados) por encima del valor medio en una dirección definida, en las coordenadas celestiales corrientes, por 11 grados de ascensión recta y seis grados Norte de latitud, es decir, aproximadamente a 15 grados Este-Suroeste de Regulus, la estrella más brillante en la constelación Leo. La velocidad del sistema solar en tal dirección se puede calcular dividiendo la máxima diferencia de temperatura, de 0,0035 grados, por la temperatura media de la radiación cósmica de fondo, 2,7 grados (el mejor valor en la actualidad), y multiplicando el resultado por la velocidad de la luz. La respuesta es una velocidad de 390 kilómetros por segundo.

Aunque esta velocidad no es mucho mayor que la de 300 kilómetros por segundo que se esperaba para el movimiento del sistema solar alrededor del centro de la galaxia, corresponde a otra dirección. Como la velocidad del sistema solar es la suma de la velocidad debida a la rotación de la galaxia más cualquier velocidad peculiar de la galaxia, podríamos a partir de nuestro valor medido y manejando correctamente los vectores, calcular la velocidad peculiar de la galaxia. Cuando hicimos esto, encontramos que la galaxia debe estar moviéndose a unos 600 kilómetros por segundo con respecto a la radiación cósmica de fondo.

Con la excepción de la variación cosinusoidal en la temperatura, se encontró que la radiación de fondo es isótropa con una exactitud superior a uno en 3000, lo que pone estrictos límites a algunos de los fenómenos que he mencionado. Si el Universo está en rotación, su velocidad de rotación debe ser inferior a 10^{-9} segundos de arco por centuria. Si existen olas gravitatorias en gran escala, no tienen energía suficiente para cerrar el Universo o invertir la expansión de Hubble convirtiéndola en implosión. La expansión misma debe ser isótropa dentro de una parte en tres mil. Tampoco hay indicación de la formación temprana de grupos de galaxias, indicando que no existió agrupamiento en escala muy grande antes del momento de desacoplo.

Quizás el más inesperado y fascinante resultado del experimento es la magnitud de la velocidad cosmológica de la galaxia que se deduce. Como el movimiento de nuestra galaxia relativo al de la de Andrómeda es pequeño (80 kilómetros por segundo), la galaxia de Andrómeda debe compartir esta alta velocidad a través del espacio. Además, se sabe que el movimiento peculiar (no de Hubble) de nuestro grupo local de galaxias relativo al más próximo grupo grande de galaxias, el de Virgo, es pequeño; por tanto, el grupo entero de Virgo debe tener una velocidad cosmológica similar a la nuestra. El cuadro que emerge es el de un gran volumen de espacio, con un radio que alcanza decenas de millones de años-luz, moviéndose a una velocidad de unos 600 kilómetros por segundo respecto al universo distante.

El cuadro se hace más complicado cuando intentamos penetrar más allá en las regiones locales de espacio. Anteriormente a nuestro trabajo, Vera C. Rubin y W. Kent Ford, Jr., del Departamento de Magnetismo Terrestre de la Institución Carnegie de Washington, habían analizado con sus colegas el movimiento de nuestra galaxia relativo a una muestra de galaxias espirales cubriendo todo el cielo a unos 100 millones de años-luz. Llegaron a la conclusión de que con respecto a esa muestra, la velocidad neta del sistema solar era de 600 kilómetros por segundo. Después de descontar la fracción de la velocidad del sistema solar que se debe a la rotación galáctica, calcularon que nuestra galaxia se está moviendo, con relación a la esfera de galaxias de referencia, a una velocidad de unos 450 kilómetros por segundo.

Nuestra medición de la velocidad peculiar de nuestra galaxia, determinada a partir de la radiación cósmica de fondo, no sólo es un tercio mayor que la velocidad de Rubin-Ford, sino que también su dirección difiere de la de aquéllos en más de 100 grados. Los dos conjuntos de mediciones de velocidad pueden ponerse de acuerdo si se admite que la esfera de galaxias de Rubin-Ford se está moviendo a una velocidad cosmológica de unos 800 kilómetros por segundo en una dirección separada unos 33 grados de aquella en que nosotros nos estamos moviendo a unos 600 kilómetros por segundo a través del "éter" de radiación que quedó como residuo de la gran explosión.

Este notable cuadro resulta todavía más sorprendente cuando uno se da cuenta de que una alta velocidad peculiar implica una velocidad todavía más alta en el pasado. A medida que una galaxia se mueve a través del espacio con una gran velocidad peculiar, llega, con el tiempo, a alcanzar a otras galaxias cuya velocidad de alejamiento corresponde a la de expansión media de Hubble. De aquí que una alta velocidad peculiar se transforme gradualmente en una velocidad típica de Hubble, con el resultado neto de que las velocidades peculiares deben disminuir con el tiempo. Extrapolando hacia atrás, se encuentra que en el momento del desacoplo, la velocidad pecular de la materia de que se formó nuestra galaxia debió ser próxima a la velocidad de la luz. Por otra parte, si la velocidad peculiar se debiera a turbulencia local o a movimiento orbital alrededor de un punto distante, tal extrapolación podría no ser correcta. La velocidad de nuestro grupo local de galaxias respecto a las galaxias próximas (a escala cósmica) de Rubin-Ford, sugiere efectivamente que existe considerable turbulencia en el Universo.

Antes de aceptar esta descripción turbulenta de la estructura en gran escala del Universo, se debería recordar que nuestra observación de la radiación cósmica de fondo muestra que, con la excepción de la componente cosinusoidal, la radiación es uniforme dentro de un error de a lo más una parte en 3000. No se ve claro cómo poner de acuerdo la naturaleza uniforme de la radiación de fondo con un alto grado de turbulencia local. Ciertamente, las velocidades peculiares locales son características del Universo actual, mientras que la radiación de fondo es una instantánea del Universo tomada hace quince mil millones de años. Es concebible que el Universo posea una estructura en gran escala, tal como la rotación de un supergrupo de galaxias, que pondrían de acuerdo los aparentemente contradictorios resultados.

Quizá la más aguda crítica del modelo de una gran explosión homogénea e isótropa sea su excesiva simplicidad para poder representar la realidad. Se suele sentir la tentación de suponer que lo desconocido es sencillo. Es posible, e incluso probable, que existan estructuras en gran escala que desempeñen un papel esencial en la determinación de la estructura del Universo. Con recientes medidas de los agrupamientos de galaxias en gran escala y de la anisotropía de la radiación cósmica de fondo, podemos quizás estar comenzando a detectar tal estructura.

Ciencia y sociedad

Fijación de nitrógeno

El elevado coste de las materias primas utilizadas en la producción de fertilizante nitrogenado sintético y el ritmo de crecimiento de la población mundial han renovado el interés en el esclarecimiento del proceso de reducción del nitrógeno molecular de la atmósfera, verificado por un número de organismos inferiores, que lo incorporan a su material proteico. En la década de los sesenta, bajo los auspicios del Programa Biológico Internacional (PBI), se hizo un gran esfuerzo investigador al respecto, y se consiguió una base sólida de datos sobre la fisiología, bioquímica, genética y ecología de los distintos sistemas fijadores de nitrógeno. La inclusión de la fijación biológica de nitrógeno entre los objetivos del PBI, en unos años en que, por la disponibilidad de fertilizante nitrogenado, se contaba con el desinterés en el tema por parte de agricultores, ecólogos y economistas, reflejaba una previsión de futuro. Hoy la fijación de nitrógeno y la actividad fotosintética y conversión de la energía solar aparecen como un binomio representativo de lo que la ciencia puede aportar a la solución de la crisis alimentaria y energética.

En la presente década surgen hallazgos de notable repercusión en el esclarecimiento del proceso. Se estudia su genética y se llega a conocer la distribución y regulación de la información que capacita a ciertos organismos vivos a fijar nitrógeno. Los genes *nif*, responsables de la codificación de aquella actividad, que se localizan vecinos al gen responsable de la histidina en el cromosoma circular de *Klebsiella pneumoniae*, son transferidos por técnicas convencionales de genética microbiana a otras mutantes no fijadoras de nitrógeno, restaurándose en ellas la actividad. Esta transferencia de genes se realiza por conjugación, transformación y transducción, permitiendo el primero transferir fragmentos relativamente grandes del genoma. Con todo, la caracterización del gen *nif* se presenta aún difícil, reconociéndose hoy unos siete determinantes del mismo, entre ellos los que determinan la incorporación de molibdeno y de hierro, metales constituyentes ambos del enzima catalizador del proceso (nitrogenasa), la incorporación de un reductor y la de una sustancia reguladora. Esta sustancia, la glutaminsintetasa, cataliza la unión del amonio procedente de la reducción del nitrógeno molecular, o de la reducción del nitrato, con el aminoácido glutamato para producir glutamina. La glutaminsintetasa se adhiere al gen *nif* iniciando la transcripción que, finalmente, conduce a la utilización de la información del *gen*. Hay que destacar que los genes determinantes de la síntesis de glutaminsintetasa se encuentran lejos de la zona del *nif*, y que el amonio verifica la regulación de su síntesis.

Las perspectivas abiertas por los estudios genéticos son amplias. La inducción de actividad nitrogenasa en *Escherichia coli* por transferencia del gen *nif* de *Klebsiella*, dada a conocer por Dixon y Postgate en 1972, a partir de experiencias realizadas en la Unidad de Fijación de Nitrógeno del Consejo de Investigaciones Agrícolas del Reino Unido, supone un hito importante. El nuevo huésped, anaerobio facultativo, sólo podrá fijar nitrógeno en ausencia de aire, condición inherente también para que esa actividad se manifieste en *Klebsiella*. Este tipo de transferencia es un avance en la idea de materializar la posibilidad de realizar análogas manipulaciones que conduzcan a la inducción de actividad nitrogenasa en organismos superiores, filogenéticamente tan distantes de las bacterias y algas azulverdosas. Dicha transferencia puede estar de hecho ocurriendo en la naturaleza, entre otras razones por el íntimo contacto de la población microbiana del suelo con los distintos huéspedes vegetales y animales, y, sin embargo, no se sabe que existan plantas ni animales que por sí mismos fijen nitrógeno. En efecto, la diferencia entre los medios de expresión de la información genética bacteriana, mecanismos de regulación, transcripción y traducción, y los correspondientes de organismos superiores, no sólo obliga a introducir el gen *nif* en el material genético de éstos sino que también se han de incorporar los medios necesarios de origen bacteriano que expresen la información introducida.

La introducción del gen *nif* en los organismos superiores, como en los cereales por ejemplo, no basta; entre otros motivos, por la ausencia en ellos de un mecanismo protector de la nitrogenasa frente al oxígeno, elemento que destruye la actividad de aquel enzima de una forma irreversible. Se necesitaría una adecuación previa del metabolismo y estructura vegetales, para lo cual los avances técnicos actuales no se hallan aún dispuestos. El desarrollo de las investigaciones en la producción de mutantes activos, más allá de las fronteras que la naturaleza nos presenta, no significa que se esté sacando el mayor partido del proceso en los sistemas biológicos ya existentes. La eficiencia de la fijación de nitrógeno en un organismo se mide en términos de cantidad de nitrógeno reducido por unidad de energía biológica (ATP) consumida. Hoy se sabe que parte del ATP consumido en el proceso se utiliza en reacciones competitivas y, entre ellas, en la producción de hidrógeno. Por supuesto que la producción parásita de hidrógeno convenientemente utilizada por el hombre para otros fines compensaría en parte el esfuerzo energético puesto por el organismo durante su actividad fijadora. La llegada de electrones al lado activo del enzima nitrogenasa va acompañada de una unión con los protones presentes para dar hidrógeno, que escapa al exterior. Ello supone una gran pérdida energética en los grandes cultivos de leguminosas, así como en la mayoría de los fijadores.

Aparte de los sistemas simbióticos tradicionales leguminosa-*Rhizobium* y no leguminosa-actinomicetes existen en la naturaleza situaciones intermedias. El descubrimiento en Australia, por Trinick en 1973, de una simbiosis entre *Trema* y *Rhizobium*, es decir, entre una no leguminosa y una bacteria que hasta la fecha se había encontrado únicamente en nódulos de leguminosas, es una muestra de las combinaciones naturales que se pueden dar. Con este hallazgo es más fácil imaginar la posibilidad de provocar sistemas simbióticos entre *Rhizobium* y otras no leguminosas cultivadas.

Un importante avance en el campo de las asociaciones simbióticas, de impacto real aún no perfilado, es la existencia de asociaciones entre bacterias fijadoras de nitrógeno y ciertas gramíneas tropicales, dada a conocer por el grupo de Johanna Döbereiner, del Consejo de Investigaciones del Brasil, concretamente entre *Azospirillum (Spirillum lipoferum)* y *Digitaria decumbens* y maíz. Extender la propiedad fijadora a las gramíneas es en extremo interesante. El factor limitante último en la fijación de nitrógeno es la fotosíntesis, limitación que se hace más patente en las leguminosas donde actúa la ruta fotosintética del carbono 3, menos eficiente que la del carbono 4 del maíz, caña de azúcar y sorgo. En éstos se aseguraría un suministro de productos fotosintetizados para la fijación biológica del nitrógeno. El avance en sí es pro-

metedor, ya que ciertamente el microorganismo puede crecer intercelularmente a las tensiones reducidas de oxígeno que imperan en las raíces de la planta sin necesitar el sistema tan complejo de la hemoglobina de los nódulos de las leguminosas; puede asimismo utilizar los productos de la fotosíntesis y el nitrógeno fijado pasar al resto de la planta por sistema vascular. No obstante, si bien se han conseguido en algunos casos aumentos en los rendimientos de cultivos de maíz estadísticamente significativos, han de buscarse otras asociaciones más eficaces. Una apropiada manipulación genética y fisiológica podría convertir a aquéllas en una asociación de mayor efectividad, y, sin duda, de enorme repercusión práctica, por cuanto supone un inmediato aumento de la producción de forraje y por ende de proteína animal.

Mención aparte merecen los intentos de obtener un modelo químico capaz de imitar la función reductora de la nitrogenasa. Con tal fin se han conseguido complejos químicos con metales de transición, capaces de reducir esa molécula de nitrógeno. La alta reactividad de los complejos de molibdeno sugiere que la nitrogenasa prefiere el molibdeno al hierro en su lado activo, pero aun cuando varios hallazgos lo indican así, no se tienen pruebas concluyentes de que los metales de transición molibdeno y hierro interaccionen con el nitrógeno molecular en su ruta hacia el amonio. En el terreno de las conjeturas se podría decir que el nitrógeno pasa a diimida (N_2H_2) coordinada al molibdeno. La disociación de este compuesto del molibdeno va seguida de una reactivación del sitio vacante por coordinación subsiguiente de la diimida con el hierro. Después de la reactivación del sitio del molibdeno, la diimida podría ser reducida a hidrazina (N_2H_4), mediando en ello una transferencia acoplada de protones y electrones. En una reactivación final la hidrazina ligada al hierro pasaría a amonio como producto final.

En los próximos diez o quince años, se podrá saber qué posibilidad existe de transferir la actividad nitrogenasa de organismos procariotas a las plantas cultivadas, y si aquélla no se da, cuáles son las razones científicas que se oponen a ello. Se podrá asimismo desarrollar el modelo químico que ayude a descifrar el funcionamiento de la nitrogenasa. Hasta es posible que se llegue a perfeccionar la asociación de bacterias fijadoras con animales, en cuyo sistema digestivo se sabe que existen.

Al margen de esto, hay que considerar que en este tipo de investigaciones, en cierto modo estratégicas, existe siempre una partida de signo negativo, ya que, en el caso que nos ocupa, se llegaría a desequilibrar el funcionamiento del ciclo biológico del nitrógeno. Sin embargo, no es menos cierto que los problemas del hombre, con su descontrolado crecimiento en número, su distribución polarizada y la limitación de los recursos explotables, hacen que el balance resulte hoy altamente positivo. Compensaciones de lo negativo serían unas investigaciones dirigidas a proveer mecanismos de desnitrificación que paliarían una eutrofización excesiva de los medios acuáticos y terrestres provocada por un futuro aumento de la actividad fijadora de nitrógeno de los organismos vivos. (C.R.B.)

Resonancias de Urano

El inesperado descubrimiento de que Urano está rodeado por cinco anillos por lo menos, semejantes a los de su vecino Saturno, pero no tan anchos, se realizó en marzo del año pasado con ayuda de un telescopio de 91 centímetros llevado a bordo del Observatorio Aerotransportado Kuiper en un vuelo sobre el Océano Indico. La finalidad original del vuelo había sido reunir información sobre la atmósfera de Urano y su forma precisa, registrando exactamente la curva de la luz de una estrella de novena magnitud (SAO 158687), conforme era ocultada por el planeta y reaparecía unos 25 minutos más tarde. Se empleó un avión porque podía volar sobre nubes celadoras y porque no era seguro que el estrecho recorrido de la ocultación intersecaría el emplazamiento de algún observatorio instalado en Tierra. Cuatro de los anillos del planeta interrumpieron la luz de la estrella tan brevemente –durante apenas un segundo– que su anchura no puede ser de más de unos 10 kilómetros. El quinto parece ser algo más ancho, quizá de 50 a 100 kilómetros. En la carta dirigida a *Nature* en la que informaron acerca de su descubrimiento, J. L. Elliot, E. Dunham y D. Mink, de la Cornell University, plantearon una pregunta fundamental: "¿Por qué son tan angostos los anillos de Urano, en comparación con los anchos anillos que rodean a Saturno?" (El brillante anillo central de Saturno tiene unos 26.000 kilómetros de anchura).

No se ha hecho esperar una teoría que explica la estrechez de los anillos de Urano. Menos de tres semanas después de abrir la carta con el descubrimiento, *Nature* recibió de S. F. Dermott y Thomas Gold, también de la Cornell, una hipótesis de trabajo. Dermott y Gold señalaron que en el caso de Saturno es preciso explicar las separaciones en los tres anillos principales y las angostas brechas en los dos anillos exteriores. La explicación generalmente aceptada es que las separaciones y brechas representan pistas a las cuales mantienen limpias o barridas de partículas las resonancias existentes entre los movimientos de las partículas existentes en los anillos y los movimientos orbitales de los mayores de entre los 10 satélites de Saturno, que giran en torno a éste por fuera de los anillos.

Invirtiendo la explicación, Dermott y Gold calculan que de entre los cinco satélites de Urano, los tres de mayor masa –Ariel, Titania y Oberón– son capaces de crear resonancias que obligarían a las partículas que circulan en torno a Urano, en cooperación con la presión de la luz solar sobre las partículas, a ocupar seis estrechas órbitas situadas a distancias que concuerdan estrechamente con las medidas durante la ocultación de SAO 158687. Dermott y Gold sugieren que el anillo ancho y más exterior no es en realidad uno, sino dos anillos muy poco distantes entre sí. Este anillo ocultó a SAO 158687 durante siete segundos antes de que la estrella desapareciera detrás del planeta, originando dos mínimos muy cercanos entre sí en la luz procedente de la estrella. Pero el anillo es peculiar, por cuanto ocultó la estrella durante sólo unos tres segundos después de que ésta hubo reaparecido desde detrás del planeta. El descubrimiento de los anillos se vio facilitado porque en el momento de la ocultación el eje de rotación de Urano estaba apuntando aproximadamente en dirección a la Tierra. Como los anillos se disponen en un plano perpendicular a dicho eje, se presentaban como una serie de anillos concéntricos.

Física de las colisiones frontales

Al igual que una colisión entre automóviles es más violenta cuando chocan de frente, las interacciones de alta energía entre partículas elementales se observan más eficientemente en aparatos en los que colisionan partículas que se muevan con igual velocidad pero en sentidos opuestos. Ya están en funcionamiento tres grandes aparatos de este tipo, y su éxito ha modificado evidentemente el diseño de la próxima generación de aceleradores. Por lo menos seis de los aparatos de alta energía planeados

o en construcción provocan colisiones entre dos haces de partículas opuestos.

La principal ventaja de las interacciones frontales es que prácticamente se libera en la colisión toda la energía empleada en acelerar las partículas. Hoy en día los protones pueden acelerarse con cierta facilidad hasta una energía de 400.000 millones de electrón voltios (400 gigaelectrón voltios, o GeV). Sin embargo, cuando una de estas partículas colisiona con un electrón en reposo, queda disponible, para la creación de nuevas partículas, una energía de menos de 28 GeV. Si dos protones, cada uno con una energía de 400 GeV, se pudieran hacer colisionar de frente, se liberaría prácticamente toda su energía (800 GeV). Para alcanzar la misma energía efectiva bombardeando un blanco fijo, se requeriría un acelerador que proporcionase una energía de unos 350.000 GeV (350 teraelectrón voltios o TeV).

Las colisiones frontales se originan colocando en anillos concéntricos o intersecantes haces de partículas que giren en sentido contrario. Una vez llenados los anillos por un acelerador auxiliar, los haces circulan de forma estacionaria durante horas. El muy bajo porcentaje de interacciones en un aparato de colisión de haces frente al correspondiente a un acelerador con blanco fijo, es el precio que hay que pagar para conseguir esta gran energía eficaz. Los haces acostumbran a pasar uno a través del otro sin interaccionar, y en los aparatos existentes sólo se produce una colisión de interés cada pocos segundos. Por esta razón son importantes el número de partículas almacenadas y el confinamiento de todas ellas a una pequeña sección transversal al objeto de optimar el rendimiento del aparato.

Dos de los grandes anillos que están ya en funcionamiento almacenan haces de electrones y de sus antipartículas, positrones, que giran en sentido contrario. Uno se halla en el Stanford Linear Accelerator Center (SLAC) y el otro está en el sincrotón de electrones alemán DESY, cerca de Hamburgo. En estos aparatos es en donde se ha observado más claramente la nueva familia de partículas con la propiedad llamada "encanto". Con el anillo del SLAC también se han obtenido resultados que sugieren la existencia de un nuevo leptón, o partícula de propiedades parecidas a las del electrón. Ambos anillos alcanzan una energía máxima de 4 GeV por haz y pronto serán reemplazados por otros aparatos mayores.

Un nuevo anillo de almacenamiento alemán, que se llamará PETRA, empleará el anillo ya existente como parte de un sistema inyector. La energía máxima disponible será de 19 GeV por haz, lo que da un total de 38 GeV. Su construcción ya se está realizando (está instalado y comprobado un octavo del anillo), y el proyecto probablemente se completará en un año. El nuevo anillo de almacenamiento de electrones y positrones de Stanford, que se llamará PEP, se está construyendo en un extremo del acelerador de electrones de dos millas del SLAC. El nuevo anillo tendrá una circunferencia de 2,5 kilómetros, aproximadamente 10 veces el tamaño del actual aparato de colisión de haces. Los planes en curso pretenden alcanzar una energía máxima de 18 GeV por haz, pero existe el propósito de aumentar este límite hasta los 24 GeV. Las colisiones tendrán lugar en seis regiones de interacción, en cinco de las cuales será posible realizar experimentos físicos. Ya han sido aprobados unos cuantos experimentos iniciales y se han construido los detectores de partículas necesarios para ellos. El propio PEP quedará terminado en 1979 o 1980; depende del flujo de las subvenciones.

Desde hace algún tiempo están operando otros anillos de almacenamiento de electrones y positrones más pequeños en los Laboratorios Frascati, cerca de Roma, y en el Instituto Serpukov de Física de Altas Energías, en las inmediaciones de Moscú. Se están construyendo también tres anillos más pequeños. En los Laboratorios Orsay, junto a París, se instalará un aparato con dos anillos tangentes. En la Universidad de Cornell se está transformando un sincrotón de electrones para que pueda almacenar haces de electrones y positrones con energías comprendidas entre 1,5 y 10 GeV. En Novosibirsk, Unión Soviética, está a punto de techarse un anillo llamado VEPP-4, cuya capacidad le permitirá alcanzar una energía de 7 GeV por haz.

Todos los aparatos precedentes almacenan electrones y positrones. El gran anillo restante que está ahora en funcionamiento almacena haces de protones. Se trata del Intersecting Storage Rings (ISR), de la Organización Europea para la Investigación Nuclear (CERN), cerca de Ginebra. Alcanza una energía máxima de 28 GeV por haz.

Se va a construir en el Laboratorio Nacional de Bookhaven, en Long Island, un aparato del mismo tipo pero mucho mayor, llamado ISABELLE. Constará de dos anillos entrelazados, cada uno de los cuales podrá mantener un haz de protones cuya energía podrá ser ajustada continuamente dentro de un intervalo comprendido entre 30 y 400 GeV. No só-

lo estas más altas energías, sino también el número de protones almacenados supondrán un mayor rendimiento frente al de los primeros aceleradores. La circulación de los protones en cada anillo equivaldrá a una corriente de unos ocho amperios. Los haces estarán confinados en los anillos y serán dirigidos por 1084 imanes superconductores instalados en un túnel de 3,7 kilómetros de circunferencia.

En el Fermi National Accelerator Laboratory (Fermilab), cerca de Chicago, que hasta ahora sólo tenía un acelerador con blanco fijo, están en estudio dos proyectos de aparatos de colisión de haces que podrían alcanzar energías incluso superiores a las de ISABELLE. Un nuevo anillo de imanes superconductores (capaz de acelerar protones hasta 1 TeV) se está instalando en el mismo túnel que el sincrotón de protones ya existente. Almacenando protones en el nuevo anillo y en el viejo podrían lograrse energías de colisión de 1,3 TeV. Otra posibilidad es almacenar simultáneamente protones y antiprotones en el anillo superconductor. En principio debería obtenerse una energía de 2 TeV, pero las técnicas para crear haces intensos y estables de protones se hallan todavía en fase de investigación.

También está en estudio, en el CERN, un anillo de almacenamiento de protones y antiprotones, entre otros varios objetivos. De éstos hay que destacar uno: se trataría del mayor sistema de almacenamiento. Sería un anillo para electrones y positrones un orden de magnitud mayor que PEP y PETRA, con una circunferencia de 40 kilómetros y una energía máxima de 200 GeV por haz.

No podemos predecir qué se hallará con estos nuevos instrumentos, pero los estudios a más baja energía han dejado dos cuestiones sin resolver a las que sin duda les será dada total prioridad. En primer lugar, el estudio de los quarks, partículas elementales que se cree constituyen la estructura del protón y de muchas otras partículas parecidas; los nuevos anillos electrón positron podrían resolver la pregunta de cuántos tipos de quarks hay. En segundo lugar, la búsqueda de las partículas llamadas *W* y *Z*, que se supone transmiten la fuerza "débil" de la desintegración beta nuclear. Quizá las partículas *W* y *Z* sean los trofeos más preciados de los físicos contemporáneos. Con masas estimadas entre los 50 y 100 GeV, probablemente estén más allá del rango de energías de PEP y PETRA, pero deben ser detectadas con ISABELLE o con los anillos de almacenamiento del Fermilab y del CERN.

El nucleosoma

La cromatina, en su nivel estructural más íntimo, está ordenada de una forma simple mediante una subestructura básica repetitiva integrada por una agrupación nucleoproteica globular discreta: el nucleosoma

L. Cornudella

La materia es el elemento fundamental del estado físico. Las continuas transformaciones experimentadas por ella, y muy especialmente su gradual organización en estructuras cada vez más complejas, nos fuerzan a introducir concepciones de naturaleza evolutiva en nuestra actitud analítica. En una amplia panorámica podemos distinguir tres grandes niveles evolutivos en este proceso de transformación cósmica. El alumbramiento del universo y su constante expansión –con sus cuerpos celestes, galaxias, nuestro propio sistema solar– regulado por mecanismos de evolución atómica. La constitución de la materia inorgánica y su posterior transformación en orgánica, con sus etapas finales dedicadas al moldeo y formación de las complejas moléculas indispensables para dar lugar a las formas más primitivas de vida, nos introduce en los dominios de la evolución química. El último nivel, que abarca la continua variación existente entre las innumerables especies animales y sus procesos estrictos de selección natural, se hace inteligible mediante los supuestos de la teoría de la evolución biológica.

Estos grandes marcos evolutivos nos permiten entrever un origen químico de la vida y hallar explicación aceptable a las transiciones que tienen lugar entre los organismos. Con todo, la cuestión fundamental de cuál es la esencia del estado vivo y sus características fundamentales resta todavía pendiente de aclaración, aunque debemos aceptar que su respuesta sea coherente con los términos evolutivos. Los lindes entre el carácter inerte y vivo de la materia se cruzan inexorablemente en el momento en que ésta adquiere la habilidad de dividirse. Ahora bien, esta división no puede ser simple en el sentido de que implique pérdida o disminución de la integridad original por fragmentación, descomposición, etcétera. Es esencial que se mantengan las características iniciales con todo su significado en el acto de la multiplicación. En realidad ha de tratarse de un proceso de auténtica reproducción.

La esencia distintiva de la materia viva ha de ser, pues, su capacidad para autoreplicarse. Esta propiedad exige la facultad de alterar la materia orgánica del medio circundante merced a un complejísimo espectro, equilibrado en su conjunto, de reacciones químicas que abarque desde el ordenado fraccionamiento de substancias exógenas hasta la posibilidad de elaboración de todos los componentes necesarios para consumar las funciones de crecimiento y desarrollo. Todo este conjunto de reacciones constituyen el conocido proceso general de metabolismo de la materia viva. En definitiva, la capacidad de autoduplicación, regulada por los cauces metabólicos, define la naturaleza íntima del estado vivo. A pesar de que los procesos de división y crecimiento podrían ser dirigidos por fuerzas meramente físicas, aunque sin alcanzar las exigencias de la fantástica complejidad de la naturaleza viva, es justificable pensar que la sabiduría evolutiva de miles de millones de años permitiese establecer, mediante una paciente selección natural, una información metabólica susceptible de ser transmitida a través de generaciones, y así lenta y pausadamente crear las condiciones singulares para la emergencia del primer organismo vivo.

La implicación quizá más trascendental de las teorías evolutivas es que no requieren el entramado de leyes especiales. La vida no sólo es una consecuencia natural de la evolución química, sino que su misma esencia es compartida por todas las formas vivientes. Es así como la célula viva adquiere el carácter de unidad fundamental constituyente de todos los organismos. La teoría celular consuma los ciclos precedentes de la evolución. Produce pasmo comprobar cómo las etapas evolutivas en las que han acaecido las mutaciones más elementales son aquellas que han consumido mayores intervalos de tiempo; tan enormes, que escapan casi a nuestra capacidad de percepción. En contraste, la sorprendente complejidad de la naturaleza viva ha alcanzado su pleno desarrollo en niveles de tiempo que pueden, por comparación, considerarse insignificantes.

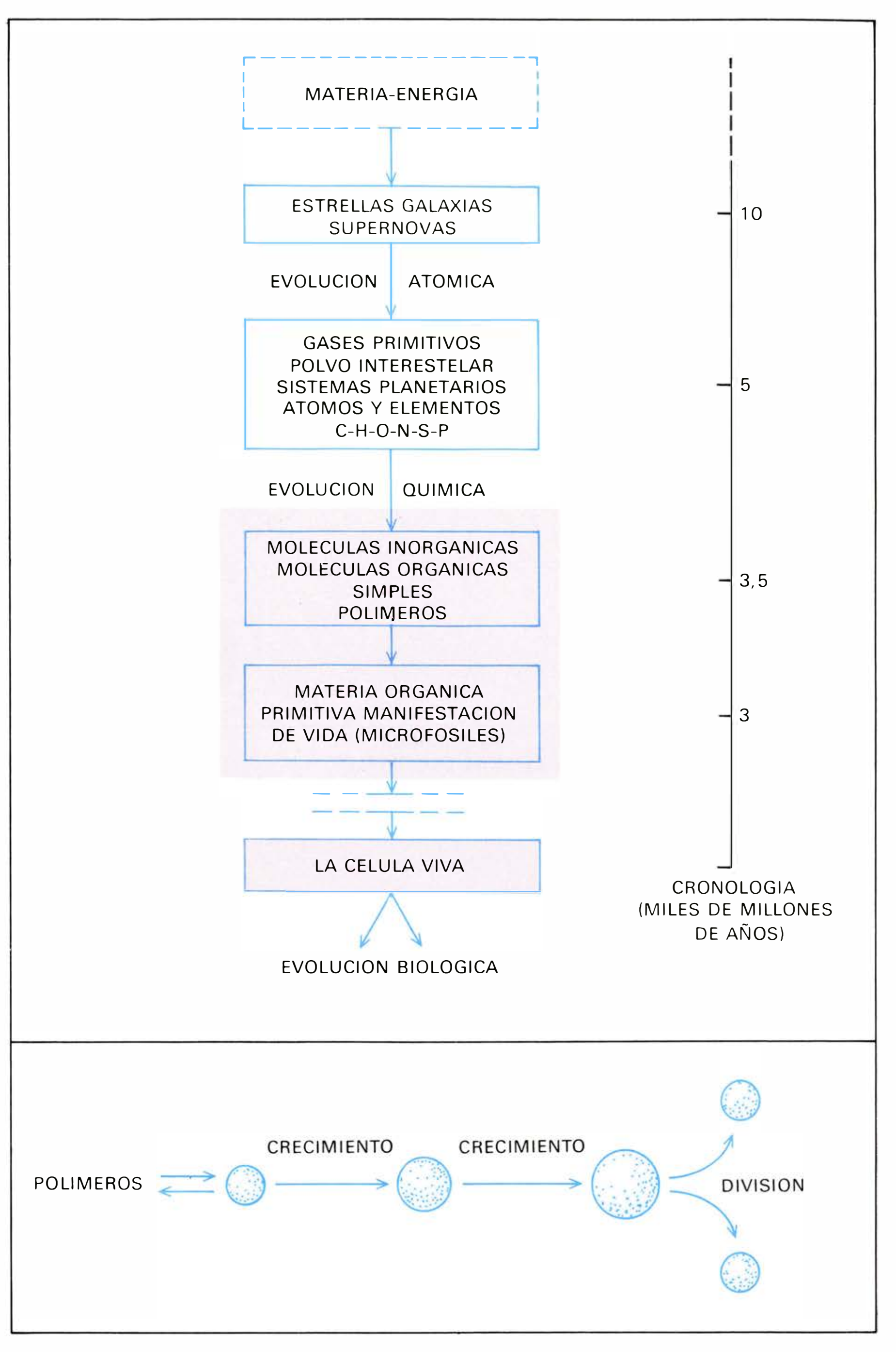
La escena ha quedado completada. Las concepciones evolutivas nos han mostrado los orígenes de la vida y conducido hasta su elemento esencial, la célula. Es importante ahora proceder a un repaso, aunque sea superficialmente, de los acontecimientos más singulares que han dado soporte a la teoría celular y abrieron el camino para desvelar los mecanismos por los cuales la célula realiza el conjunto de reacciones que sustentan los procesos vitales. Ello permitirá a continuación abordar con conocimiento de causa el estudio de su organización estructural y de su función.

Apenas hace un siglo y medio, dos naturalistas –Schleiden y Schwann– propusieron al unísono el fundamento de la teoría celular: los tejidos de todas las plantas y animales están constituidos por pequeñas unidades básicas llamadas células. De mayor trascendencia si cabe fue su aseveración de que todas las células derivan de otras predecesoras mediante un proceso de división celular. Las células se desarrollan y en un instante determinado dan lugar a dos células hijas, las cuales, debido a la división del núcleo de su progenitora, adquieren cada una de ellas un nuevo núcleo. Este componente celular posee un número fijo de cromosomas. Inmediatamente antes de la división celular propiamente dicha, estos cuerpos experimentan un proceso de partición para dar dos mitades que conservan íntegramente los caracteres originales. Gracias a este mecanismo, el núcleo celular dobla el número de sus cromosomas y lo distribuye equitativamente entre las células hijas de

manera que su complemento cromosómico final permanezca idéntico al de partida.

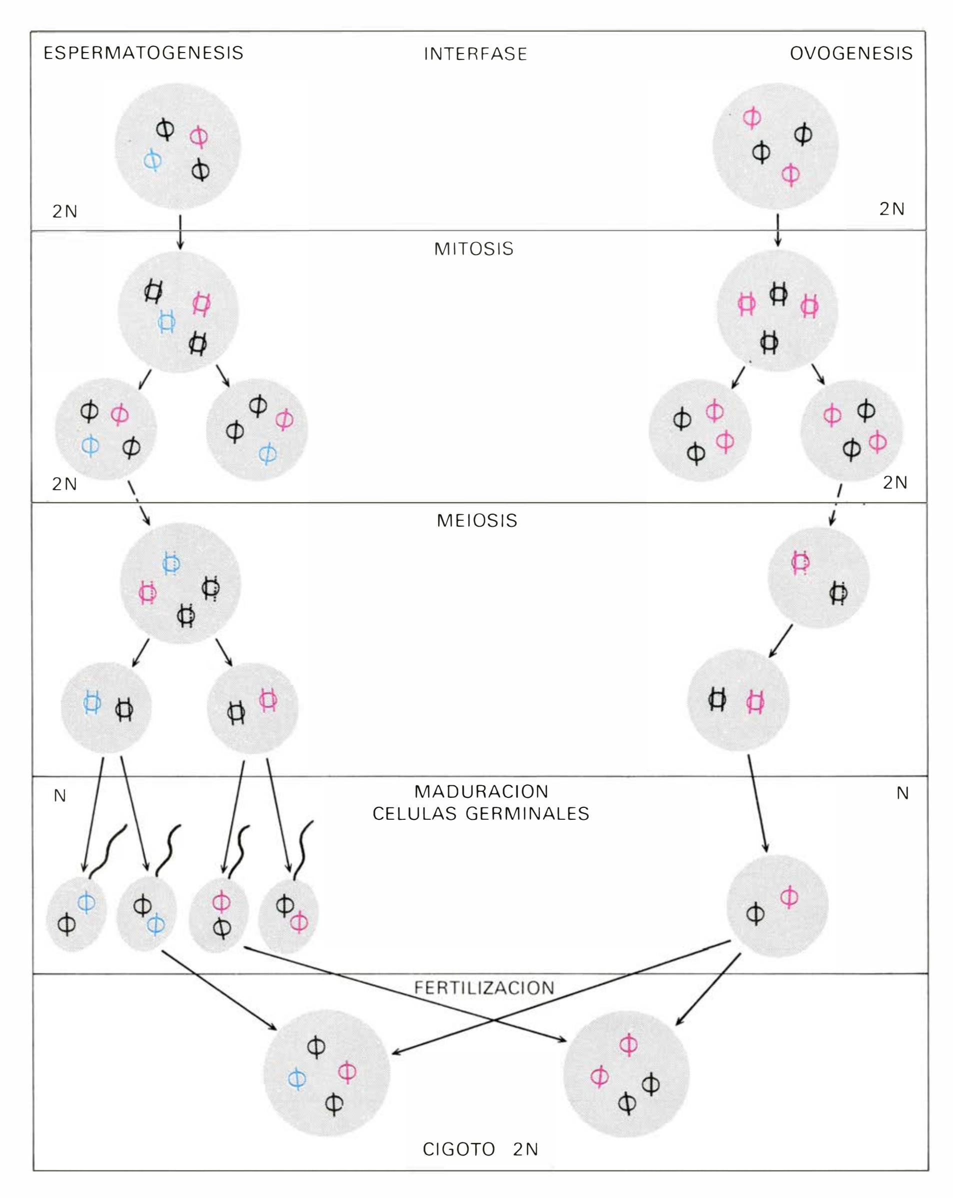
La manera de desarrollarse este proceso de multiplicación celular que hemos visto muy esquemáticamente, y por el cual las características de los cromosomas se van perpetuando, sugiere la existencia de una correlación íntima con la capacidad de transmisión de los caracteres hereditarios de generación en generación. La prueba experimental demostrativa de la relación entre la duplicación cromosómica y la herencia fue aportada hacia finales del siglo pasado por Mendel al establecer, basándose en sus observaciones, los principios hereditarios básicos. La transferencia de los diversos rasgos distintivos de un individuo es controlada por pares de factores cromosómicos. La causa de las conocidas dificultades con que tropezó Mendel para conseguir pleno reconocimiento de sus hallazgos fue probablemente debida a la ausencia en aquellos momentos de una prueba directa en favor de la relación entre el comportamiento cromosómico durante la división celular y el fenómeno de la transmisión hereditaria. Fue Sutton a principios de este siglo quien postuló la integración del gen en el cromosoma, asentando de esta manera las bases de la teoría cromosómica de la herencia.

Las reglas hereditarias de Mendel y los postulados de la teoría cromosómica de la herencia constituyeron el punto de arranque de uno de los esfuerzos más intensos realizados por el hombre para esclarecer la identidad química del gen. El comienzo de estos estudios fue lento y los resultados y avances al principio resultaron parcos. La razón de ello la podemos hallar en la singular habilidad del gen de autoduplicarse. Esta característica planteaba graves problemas de naturaleza química, pues exigía la intervención de mecanismos estructurales, hasta entonces insospechados, que facilitasen la copia fiel de moléculas de gran complejidad para dar réplicas precisas de sí mismas. Durante largo tiempo, el conocimiento de la organización del gen y su capacidad química para regular los caracteres celulares permaneció inaccesible. Lentamente, con todo, los hallazgos se sucedieron. Primero fueron los aminoácidos y su facultad de configurarse en largas cadenas poliméricas constituyendo las proteínas. A continuación se produjo el descubrimiento de la existencia de los ácidos nucleicos, llamados así por haber sido detectados inicialmente en el núcleo celular, ADN (ácido desoxirribonucleico) y ARN (ácido ribonucleico). En un principio estas dos macromoléculas se consideraron como candidatos pobres para serles atribuida una naturaleza genética, puesto que su composición no exhibía la exuberante diversidad de las moléculas proteicas.



LA CRONOLOGIA EVOLUTIVA permite fijar cada vez con mayor precisión los acontecimientos trascendentales de la historia del universo. A lo largo del curso evolutivo la masa ígnea original de materia inerte se transforma en materia viva estableciendo una dependencia directa entre los niveles físicos y biológicos. La teoría del origen químico de la vida nos sugiere cómo ésta y su capacidad intrínseca de dividirse y transmitirse resultan una consecuencia natural de la evolución.

La primera prueba concluyente acerca de la especificidad genética del ADN fue aportada por Avery en 1944 al conseguir alterar los caracteres hereditarios de una cepa bacteriana mediante la administración de ADN procedente de un origen totalmente distinto. Esta trascendental observación motivó un decidido impulso para nuevos estudios en el campo de la química de los ácidos nucleicos. Las contribuciones más significativas fueron debidas a Erwin Chargaff. El grupo de la universidad americana de Columbia, encabezado por Chargaff, consiguió establecer las razones relativas de las cuatro bases nitrogenadas del ADN. Este fulgurante avance en el conocimiento de la composición química de esta substancia permitió a Watson y Crick dilucidar el enigma de la conformación estructural



TRANSMISION GENETICA en los procesos de espermatogénesis y ovogénesis en el caso de dos células germinales diploides, conteniendo cada una un par de autosomas homólogos *(negro)* y de la dotación de cromosomas sexuales X *(rojo)* e Y *(azul)*. Las células primordiales, por mitosis sucesivas, originan células hijas idénticas. En la fase siguiente de meiosis el número de cromosomas es reducido y en la subsiguiente fertilización queda restituida la composición diploide normal de cromosomas.

de esta macromolécula biológica: una doble hélice complementaria. La importancia de esta organización estructural resultó de una trascendencia incalculable. El modelo de la doble hélice, con las dos hebras enlazadas entre sí, era capaz de explicar el enigma fundamental de la duplicación genética. La hélice satisfacía la condición de que sus hebras pudieran separarse al girar la molécula sobre su eje sin que fuese necesaria la producción de incisiones en su estructura. Una vez separadas las cadenas, éstas podían actuar de moldes para dar lugar, mediante mecanismos enzimáticos, a la síntesis complementaria de sus réplicas.

Las células de la mayoría de los organismos superiores contienen ADN en una proporción muy superior a las de los organismos procariotas. La casi totalidad de este ADN se halla localizado en una fracción subcelular –el núcleo– protegida por membranas que la diferencian del resto de fracciones y cuerpos celulares. El núcleo contiene además una gran cantidad de proteínas que se hallan involucradas con distinto grado de participación en los mecanismos de síntesis, proceso y elaboración de moléculas de ARN. La molécula dúplex de ADN en una célula típica se halla directamente asociada con un solo grupo específico de proteínas nucleares –las histonas– constituyendo el material genético de naturaleza nucleoproteica. En este complemento de histonas se distinguen cinco clases de proteínas y su peso global equivale al de la molécula de ácido nucleico.

El conjunto del material genético propiamente dicho que se halla empaquetado en el interior del núcleo recibe el nombre de cromatina y, como ya hemos indicado, se distribuye entre sus cromosomas definiendo los caracteres de la célula, o sea, constituyendo el genoma distintivo celular. Una característica específica de los organismos superiores es que sólo una pequeña parte del genoma expresa su potencial genético mientras que el resto permanece inactivo –sin expresarse– a lo largo del ciclo vital de la célula. Este hecho es claramente indicativo de un elevado nivel de complejidad en la organización estructural de la cromatina. Buena prueba de ello la tenemos en el grado de compactación que confieren las histonas al interaccionar con el ADN. La cadena de ADN que se halla en el interior de un núcleo al extenderse alcanzaría la longitud de un metro; ahora bien, a consecuencia de su asociación con las histonas, el complejo nucleoproteico resultante se pliega sobre sí mismo de tal forma que puede contenerse con facilidad en un cuerpo celular de una micra de diámetro, que corresponde al orden dimensional del núcleo.

A pesar de que es obvia la existencia de una íntima y directa dependencia entre la organización estructural del material nucleoproteico y su función, en la actualidad todavía es poco lo que se sabe sobre cómo la cromatina se halla estructurada y cómo a través de su organización ejerce una influencia decisiva en los procesos de expresión genética. Sabemos que sólo una pequeña proporción del ADN celular está integrado en genes implicados en la codificación de proteínas, es decir, transcribiendo activamente, mientras que los genes restantes permanecen reprimidos, sin expresarse. Esta situación a primera vista resulta incongruente con los postulados de la economía biológica, pues lleva a pensar que una porción grande del genoma resta inactivo careciendo de finalidad alguna. Los conocimientos actuales sugieren que estas regiones aunque no participen en el proceso de transcripción no son superfluas sino al contrario, tienen asignado un cometido de control de los mecanismos de expresión y simultáneamente intervienen en la propia organización del cromosoma. Estas funciones claramente son indispensables y de gran trascendencia para una correcta regulación de la actividad genética.

Los cauces por los que transcurren estos procesos biológicos nos descubren una serie de aspectos hoy todavía problemáticos y que aguardan una correcta resolución. La naturaleza íntima de la asociación entre el ADN y las histonas

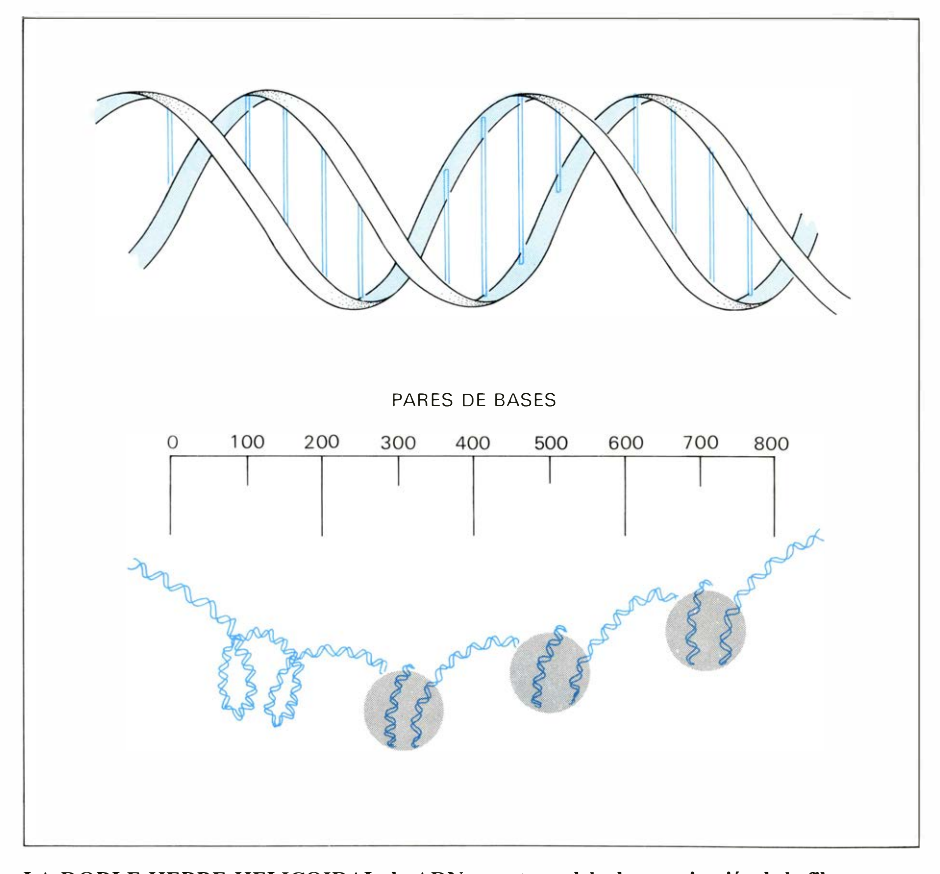
es uno de ellos. El fenómeno de la condensación cromosómica que tiene lugar a lo largo del ciclo celular es consecuencia de la interacción de los dos componentes cromatínicos mencionados, pero carecemos de datos aclaratorios acerca de la naturaleza y características de la señal celular que determina su inducción y de cómo aquella interacción incide sobre los mecanismos de regulación genética. Otra laguna importante la encontramos en el tipo de los complejos de histonas que constriñen al ADN. Las formas de activación y control genéticos en los tejidos celulares diferenciados no han sido aún descifrados. Los elementos distintivos entre las regiones genéticamente activas e inactivas del ADN continúan siendo un enigma, así como las transiciones que exhibe la cromatina a nivel estructural y su repercusión en el desarrollo celular. Este conjunto de temas nos delimita una problemática en cuya solución reside el conocimiento de la esencia de los fenómenos biológicos que son el testimonio de la dinámica del estado vivo. Veamos ahora qué cota de este conocimiento hemos conseguido alcanzar y consolidar.

Es evidente que un tema de la envergadura de la organización del genoma, con la trascendencia singular de sus funciones, haya sido abordado desde ángulos distintos y con enfoques experimentales diferentes. Examinemos las aportaciones más descollantes de los distintos campos científicos al esclarecimiento de esta cuestión. Los resultados obtenidos mediante la aplicación de la poderosa técnica analítica de la difracción de rayos X indican que la estructura básica de la cromatina se caracteriza por exhibir un diagrama de difracción en el que aparece una serie regular de anillos de difracción. Esta regularidad en la difracción sugiere la existencia de una ordenación estructural de tipo repetitivo en la fibra nativa de cromatina. La eliminación por medios químicos de las histonas asociadas con el ADN conlleva la desaparición del diagrama de difracción típico. Estos hechos conocidos ya desde hace algún tiempo permitieron a Wilkins proponer un modelo de estructura en el cual las histonas envolvían a la molécula de ADN, determinando el arrollamiento sobre sí misma y la adopción de una conformación en superhélice.

En apoyo de los resultados e interpretaciones de los diagramas obtenidos por rayos X podemos considerar los análisis por difracción de neutrones. Esta técnica posee un fundamento físico similar a la técnica de los rayos X, pero exhibe mayor poder de resolución. Esta circunstancia favorable se debe a la posibilidad de distinguir entre la cooperación de cada componente al diagrama de difracción. Las muestras de cromatina que se emplean en esta clase de análisis se preparan en forma de fibras que contengan ADN y sus proteínas asociadas. La difracción debida a los átomos de hidrógeno frente a los neutrones es anómala y de signo negativo. Este hecho permite diferenciar al hidrógeno del resto de átomos cuyos valores de difracción suelen ser positivos. Las proteínas poseen, en general, una mayor proporción de átomos de hidrógeno que la molécula de ADN; en consecuencia, los diagramas de difracción reflejan de distinto modo el comportamiento del ADN y del componente proteico en las fibras de cromatina. El superior refinamiento de la difracción de neutrones ha permitido consolidar los resultados debidos a los rayos X y precisar la participación de cada componente de la cromatina en los diagramas de difracción.

El fundamento físico de las técnicas de difracción determina que la interpretación de sus resultados sea particularmente árida y de difícil transposición a formas de diseño concretas. Debido a esta situación, es evidente que la microscopía electrónica, merced a su capacidad de visualizar directamente los materiales objeto de estudio, aportase su colaboración a dilucidar los problemas planteados. Sin embargo, los procesos preparativos inducen un determinado grado de desnaturalización en las muestras que limita la significación de sus resultados. Hecha esta advertencia, podemos indicar que la aplicación de la microscopía electrónica al examen de la cromatina permitió detectar en un principio la presencia de unos cuerpos esféricos distribuidos a lo largo del material fibroso y conectados entre sí por filamentos de menor grosor aparentemente desprovistos de formas particuladas. La existencia de una organización regular sugerida por los diagramas de difracción es totalmente compatible con la presencia de unas partículas ordenadamente distribuidas, ya que este ordenamiento difractaría, de acuerdo con los diagramas que se obtienen en la realidad. El conjunto de datos experimentales que acabamos de comentar nos suministra un conocimiento estructural pero no suficiente, ya que nada nos dice acerca de la funcionalidad de estas estructuras y de su propia e íntima composición.

La investigación química en este campo ha conseguido tras repetidas tentativas solventar con éxito la extracción del complemento de histonas de la croma-



LA DOBLE HEBRE HELICOIDAL de ADN en este modelo de organización de la fibra cromatínica se arrolla en el espacio para constituir una superestructura que facilita la asociación con el complejo globular de histonas, determinando la formación de subunidades estructurales distribuidas ordenadamente a lo largo de la fibra. Entre los puntos medios de la hebra de ADN que interconecta partículas nucleoproteicas adyacentes existe una periodicidad de 200 pares de bases.

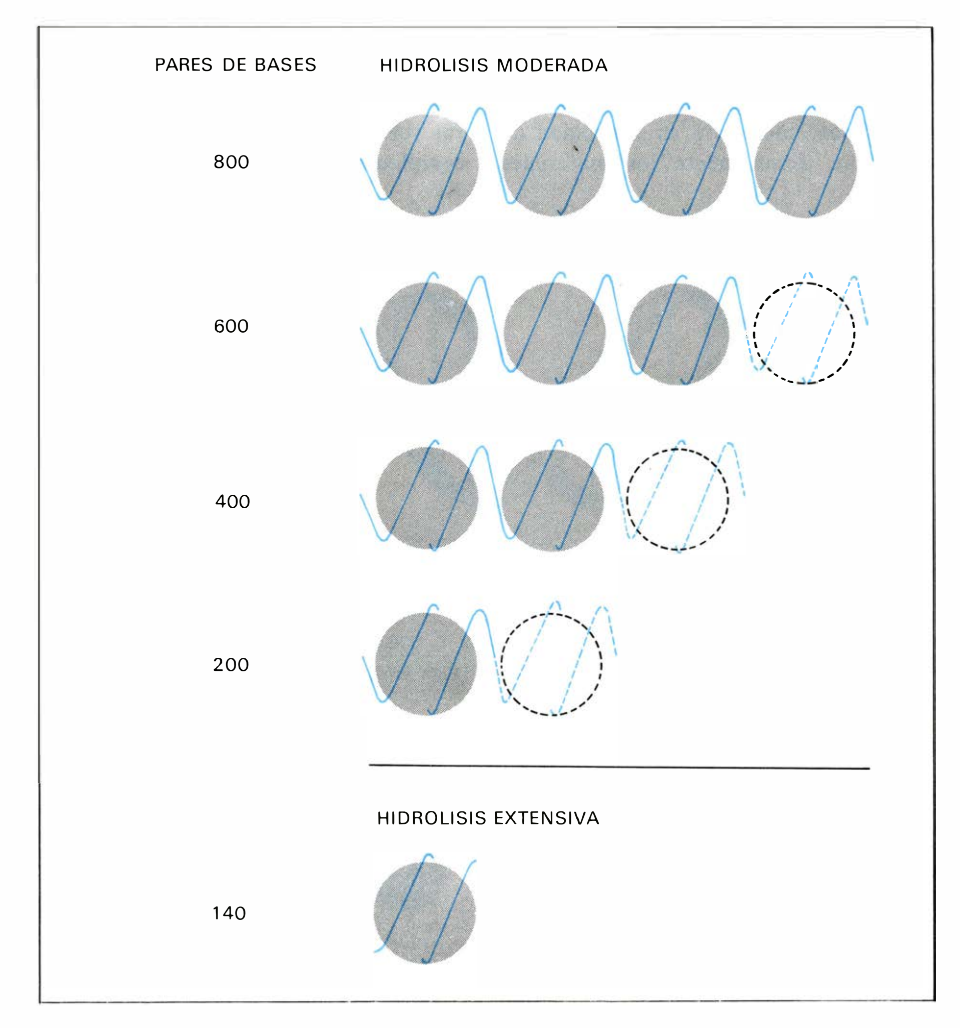
tina mediante el empleo de métodos extractivos suaves, soslayando así el problema de la fácil tendencia a la agregación que presentan estas proteínas. El fraccionamiento de esta dotación de proteínas con resinas de intercambio ha conducido a la obtención y aislamiento de tres agrupaciones proteicas discretas. Una de ellas está constituida por un tetrámero que contiene dos moléculas de cada una de las histonas pertenecientes a la clase denominada rica en el aminoácido arginina. Las dos proteínas del grupo de las histonas parcialmente enriquecidas en lisina, en estas condiciones de fraccionamiento, constituyen un complejo de naturaleza oligomérica. Finalmente, la histona con una mayor proporción del aminoácido lisina aparece separada al fraccionar sin hallarse combinada con ninguna de las restantes histonas. Estos resultados que acabamos de describir han podido ser reproducidos por métodos químicos distintos, como el de la formación de oligómeros mediante la utilización de agentes intercalantes.

Una pieza de información de gran significado lo constituye el hecho de que los distintos oligómeros formados por las histonas, al ser reasociados al ADN, recuperan la forma original de difracción de la propia cromatina nativa. Esta última circunstancia constituye una prueba directa de que los distintos complejos de histonas obtenidos por medios químicos no son consecuencia de artefactos experimentales, sino que corresponden a una presencia real en la cromatina.

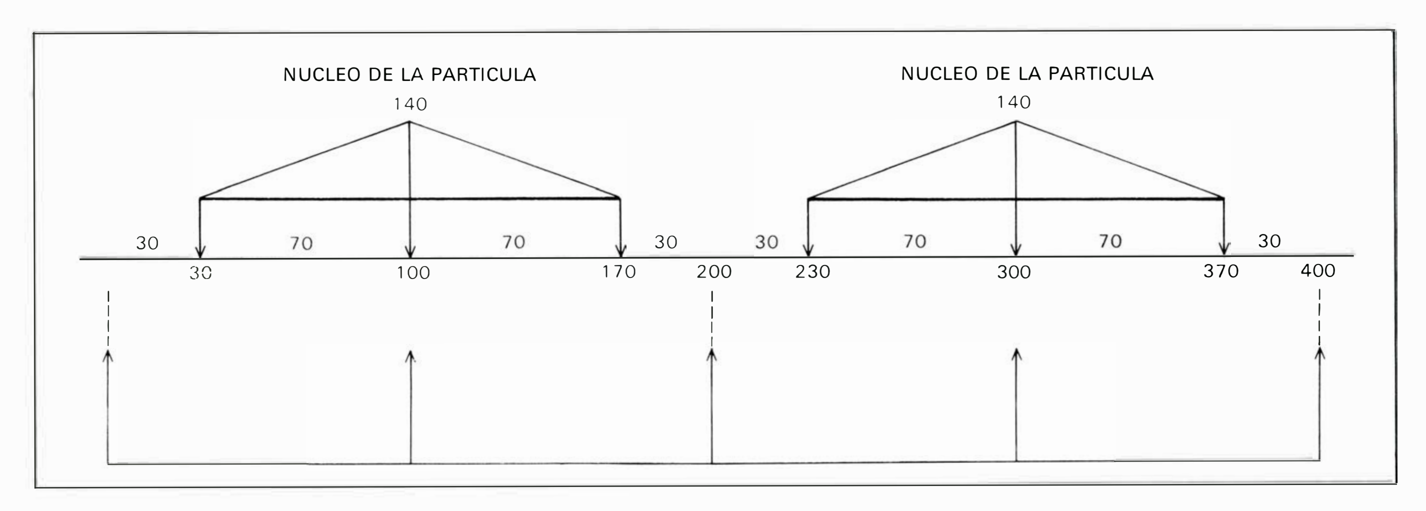
Hemos dicho al comienzo que las histonas y el ADN se hallan en la cromatina en cantidades ponderales semejantes. Si asumimos ahora que el tetrámero formado por las histonas ricas en arginina se halla asociado a dos copias del dímero debido a las dos proteínas del grupo de las parcialmente enriquecidas en lisina, ambas agrupaciones constituirán un complejo octamérico. Si ahora calculamos el peso de ADN equivalente al de estas ocho moléculas de histona comprobaremos que cualquier tipo de interacción entre ambos componentes debe involucrar un segmento de la doble hebra del ácido nucleico correspondiente a 200 pares de bases nitrogenadas. Podemos imaginar así una partícula nucleoproteica compuesta por el octámero de histonas y su correspondiente segmento de ADN. Las dimensiones que se deducen para una partícula de las características señaladas nos llevan a averiguar con gran sorpresa que ostentan una muy precisa correspondencia con las dimensiones que deben asignarse al elemento estructural regular sugerido por los análisis de difracción, para que se cumplan los diagramas experimentales. Esta similitud se hace extensiva también a los valores calculados para las formas esféricas que aparecen en las micrografías electrónicas.

De la contemplación conjunta de los resultados obtenidos por técnicas tan diversas, se va perfilando cada vez más nítidamente un modelo de conformación estructural para la cromatina. La bien establecida asociación entre las histonas y el ADN es perfectamente realizable mediante una agrupación fundamental integrada por una unidad proteica multimérica, como la propuesta por los estudios químicos, que interacciona con fragmentos específicos de ADN de longitud determinada. Esta agrupación básica que podemos calificar de auténtica subestructura, de acuerdo con la prueba de los análisis de difracción, debe distribuirse a lo largo de la fibra cromatínica de manera regular y repetitiva. Este ordenamiento generaría en definitiva una macroestructura de gran flexibilidad, capaz de doblarse y arrollarse sobre sí misma para satisfacer los requisitos de condensación y empaquetamiento exigidos por el tamaño real del cromosoma.

El atractivo del modelo estructural que acabamos de comentar ha hecho redoblar los esfuerzos de la investigación para lograr su plena confirmación. Las pruebas de orden físico han cumplido con su cometido y nos han revelado unas formas estructurales, aunque sin alcanzar a indicarnos la significación de las mismas. Mediante los planteamientos químicos se ha comenzado a descifrar la composición de estas estructuras y avanzar conocimientos sobre su función; de todas maneras, su alcance se ha visto limitado al polarizarse los intentos hacia el estudio de los componentes cromatínicos por separado, o sea, fuera de su contexto natural. La detección de los



ORGANIZACION DE LA CROMATINA mediante una subestructura nucleoproteica discreta (nucleosoma); determina que el segmento de doble hebra de ADN *(azul)* asociado a la agrupación globular proteica *(sombreado)* sea resistente a la acción nucleolítica enzimática. Los tramos de hebra internucleosomal quedan menos protegidos del ataque enzimático y facilitan, en condiciones hidrolíticas moderadas, la formación de oligómeros y multímeros del nucleosoma. Estas subunidades multiméricas son accesibles a la nucleasa dando lugar a la partícula central del nucleosoma.



EXTENSION de la doble hélice de ADN integrada en la fibra de cromatina; permite apreciar los puntos de acceso de las diversas nucleasas. La periodicidad exhibida por la nucleasa micrococal produce nucleosomas constituidos por un núcleo central resistente de 140 pares de bases que son digeridos para dar lugar al núcleo de la subunidad. La DNasa II actúa de forma similar pero con una frecuencia de 100 pares de bases, indicativa de que el nucleosoma no es homogéneo sino que puede desdoblarse en dos mitades. La sensibilidad a la DNasa I es mucho mayor.

complejos oligoméricos de las histonas que hemos descrito, no constituyen con todo una prueba concluyente de su existencia en la cromatina nativa.

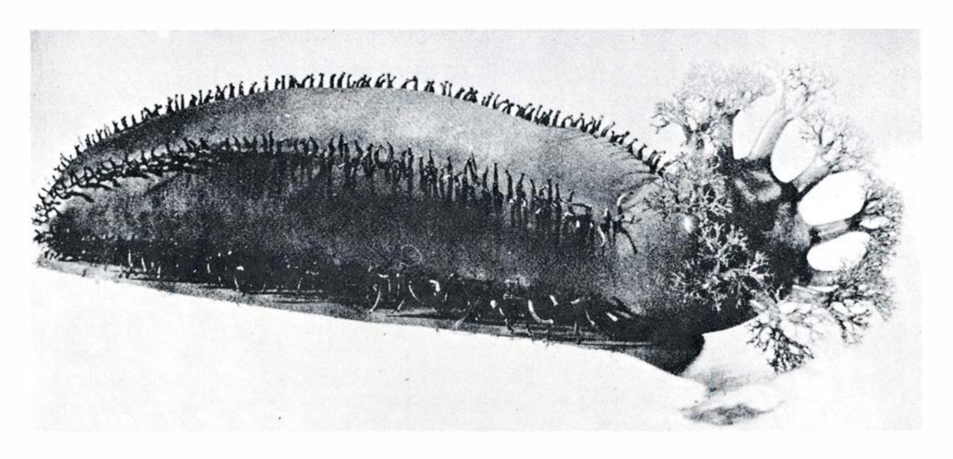
Hace escasamente tres años que, conscientes de esta situación en cierto modo precaria, se intensificaron las investigaciones y, lo que es más importante, emergieron nuevas ideas para abordar el problema desde una perspectiva genuinamente bioquímica, es decir, partiendo del supuesto de que la cromatina es un todo integral no susceptible de ser estudiado fraccionadamente por conllevar el deterioro de su integridad funcional. Como es previsible tratándose de planteamientos bioquímicos, los enzimas aparecieron como la herramienta experimental de mayor eficacia.

La primera observación realizada fue casi debida al azar. Se observó que al incubar *in vitro* extractos nucleares de hígado de rata en presencia de iones calcio y magnesio, la cromatina sufría un proceso de digestión muy peculiar y enteramente reproducible. Las características de la reacción hicieron sospechar una causa de origen enzimático, como así se confirmó poco después. Se trataba de un enzima de acción nucleolítica de carácter endógeno, perteneciente al propio tejido celular, que se activaba por la presencia de los dos iones metálicos divalentes mencionados. Al proceder a la extracción y purificación del ADN de los digeridos, y someterlo a análisis electroforético, se pudo comprobar que se obtenían una serie de bandas electroforéticas indicativas de que el ácido nucleico, por la acción hidrolítica del enzima, se segmentaba con una periodicidad predeterminada.

Este hecho ha provocado un auténtico alud de experimentos diversos en los que cromatinas preparadas a partir de gran variedad de especies celulares distintas son sometidas a reconocimiento por enzimas del tipo de las nucleasas, destacando como más empleadas la nucleasa de origen micrococal o estafilocócica y las DNasas I y II. Se ha podido comprobar que estas moléculas enzimáticas actúan fragmentando sistemáticamente el ADN de forma regular. Veamos con cierto detalle cómo transcurre la acción enzimática. Si la cromatina se incuba brevemente en presencia de nucleasa, ésta produce unos cortes en la fibra que dan lugar a una serie de partículas nucleoproteicas que contienen fragmentos de ADN de peso molecular determinado. El hecho más notable es que los segmentos de ácido nucleico de las partículas de mayor tamaño son múltiplos integrales del correspondiente a la partícula más pequeña que se logra durante la digestión enzimática. Al tomar ahora el producto total de la hidrólisis y someterlo a análisis de sedimentación por gradiente en la ultracentrífuga, se obtienen unos diagramas que muestran una serie de picos de absorción correspondientes a las partículas agrupadas por sus tamaños. Esta experiencia puede realizarse en escala preparativa, permitiendo separar los distintos grupos de partículas y proceder a su posterior análisis y caracterización individualmente. La partícula nucleoproteica menor que resulta de la digestión posee un fragmento de la hebra de ADN cuya longitud equivale a 200 pares de bases nitrogenadas, o sea, a una doble cadena de 200 desoxirribonucleótidos consecutivos. Como hemos dicho, las partículas mayores contienen trozos de ácido nucleico que son múltiplos del más pequeño.

El resultado de la acción nucleolítica en la cromatina es tan evidente que, sin dificultad, puede deducirse que la fibra está organizada mediante una distribución regular y periódica de una partícula unitaria que posee un segmento de ADN de longitud precisa, dando así lugar a una subunidad monomérica. Las agrupaciones mayores se deben a series sucesivas de dos o más monómeros cuya integridad ha sido respetada por el enzima, constituyendo de esta forma dímeros, trímeros, tetrámeros y multímeros en general. Este comportamiento de la cromatina frente a la nucleasa micrococal indica que la fibra comprende una subestructura que se repite regularmente con un determinado espaciado y que ofrece unas zonas específicas de reconocimiento que son accesibles a la molécula de enzima y por donde ésta produce las incisiones que dan lugar a las distintas agrupaciones nucleoproteicas.

Cuando la cromatina se somete a una acción más enérgica de la nucleasa de estafilococo, el resultado de la digestión es muy parecido al que se acaba de describir. Solamente es posible distinguir una particularidad. En condiciones hidrolíticas más drásticas aparece una partícula de tamaño algo menor al del monómero. El análisis electroforético nos señala de nuevo que el segmento de ADN correspondiente posee una longitud equivalente a 140 pares de bases nitrogenadas. Esta observación nos lleva a suponer que, a medida que la reacción enzimática progresa, la nucleasa, además de reconocer las regiones de espaciado entre las partículas monoméricas, inicia el proceso degradativo sobre la propia estructura unitaria liberando parte, o mejor dicho, recortando el fragmento de la hebra de ADN que contiene. Esta



LOS HOLOTURIOIDEOS (cohombros de mar) son una clase de equinodermos muy abundantes, con una amplia distribución en los fondos costeros. La fecundación externa y la deposición de las células germinales en el agua hace que su aparato reproductor adquiera considerables proporciones.

circunstancia sugiere que no todo el ácido nucleico que integra el monómero ofrece el mismo grado de accesibilidad al enzima. Sobre este aspecto concreto volveremos a insistir.

Al tratar de la acción de las nucleasas y particularmente de la de origen micrococal sobre la cromatina hemos polarizado el comentario hacia las vicisitudes experimentadas por el ADN sin dar referencia alguna de lo que acontece con el complemento de proteínas. Se ha hablado de la posibilidad de separar los productos de la digestión enzimática por sedimentación en la ultracentrífuga preparativa y su posterior exploración analítica. Evidentemente, el análisis químico no se ha limitado a indagar sobre el ADN sino que ha cuidado también de examinar el constituyente proteínico de las partículas. La subunidad monomérica contiene todas las clases de histonas con excepción de la histona H1, o sea, la más rica en el aminoácido lisina. El resto de agrupaciones de mayor tamaño exhibe sin salvedad alguna las cinco clases de histonas conocidas. Si traemos a colación ahora el dictamen analítico acerca de la dotación de histonas de la cromatina, nos daremos inmediata cuenta de que el octámero resultante contenía el conjunto de las histonas con excepción de la H1 y además, fundados en el elemento ponderal, se deducía la correspondencia con un segmento de ADN equivalente a 200 pares de bases.

La distribución regular puesta en evidencia por la difracción de rayos X adquiere la forma de partículas esféricas en las micrografías electrónicas. Estos corpúsculos contienen ADN e histonas en la proporción de 200 pares de bases por 8 moléculas de proteína, exceptuada la H1. Las nucleasas inciden en los espacios libres entre partículas adyacentes o grupos de ellas liberando monómeros y sus múltiplos integrales. Con toda lógica hay que asignarle un cometido a la histona H1 en este cuadro. Parece muy probable que esta proteína ostente una función de tipo estructural, o sea, actúe de fuerza cohesiva entre las partículas a modo de una grapa que las mantiene unidas. Cabe también la hipótesis de que la H1 esté implicada a un nivel de superestructura, condicionando los plegamientos de la fibra cromatínica. Este comportamiento justificaría el por qué en las digestiones enzimáticas, particularmente cuando no son enérgicas, se obtienen agrupaciones de tamaños variables. Imaginemos el enzima cortando en el filamento de ADN entre partículas. En las zonas en las que la histona H1 mantiene aglutinadas varias subunidades, el acceso resultaría dificultado y en consecuencia, tras omitirlas, procedería hacia regiones de la fibra cromatínica menos protegidas. El efecto inmediato sería la producción de una agrupación oligomérica en el tramo de hebra no afectado. Sobre este particular debe indicarse que experiencias realizadas en nuestro laboratorio demuestran que la eliminación selectiva de la histona H1 de la cromatina, sin alterar los productos de la digestión, transforma a ésta en un substrato mucho más sensible a la hidrólisis enzimática.

Especial relieve exhibe la circunstancia de que las partículas nucleoproteicas que se obtienen en las reacciones enzimáticas por sí mismas son resistentes al ataque de los enzimas. Esta conducta indica que el fragmento de ADN correspondiente se halla protegido por las histonas presentes de manera que resulta inasequible a la nucleasa. Recientes mediciones hidrodinámicas de las partículas ponen de manifiesto que su comportamiento físico corresponde al de estructuras globulares discretas. Estamos comprobando cómo todas las pruebas experimentales, por distinto que sea el planteamiento teórico del que procedan, apuntan hacia un mismo sentido. La cromatina en su nivel estructural más íntimo está ordenada de una forma simple mediante una subestructura fundamental repetitiva integrada por una agrupación nucleoproteica globular discreta, que contiene histonas asociadas a un fragmento de doble hebra de ADN de longitud determinada. Esta subunidad se repite con periodicidad uniforme y aparece interconectada por filamentos de ácido nucleico carentes de la protección de proteínas y, en consecuencia, constituyendo puntos de reconocimiento sensibles al ataque de las nucleasas. Una imagen sencilla de este ordenamiento de la cromatina la tenemos en las cuentas de un collar. Esta estructura unitaria fundamental ha sido denominada nucleosoma por analogía con otros corpúsculos y orgánulos intracelulares (polisomas, ribosomas, lisosomas, peroxisomas, etcétera).

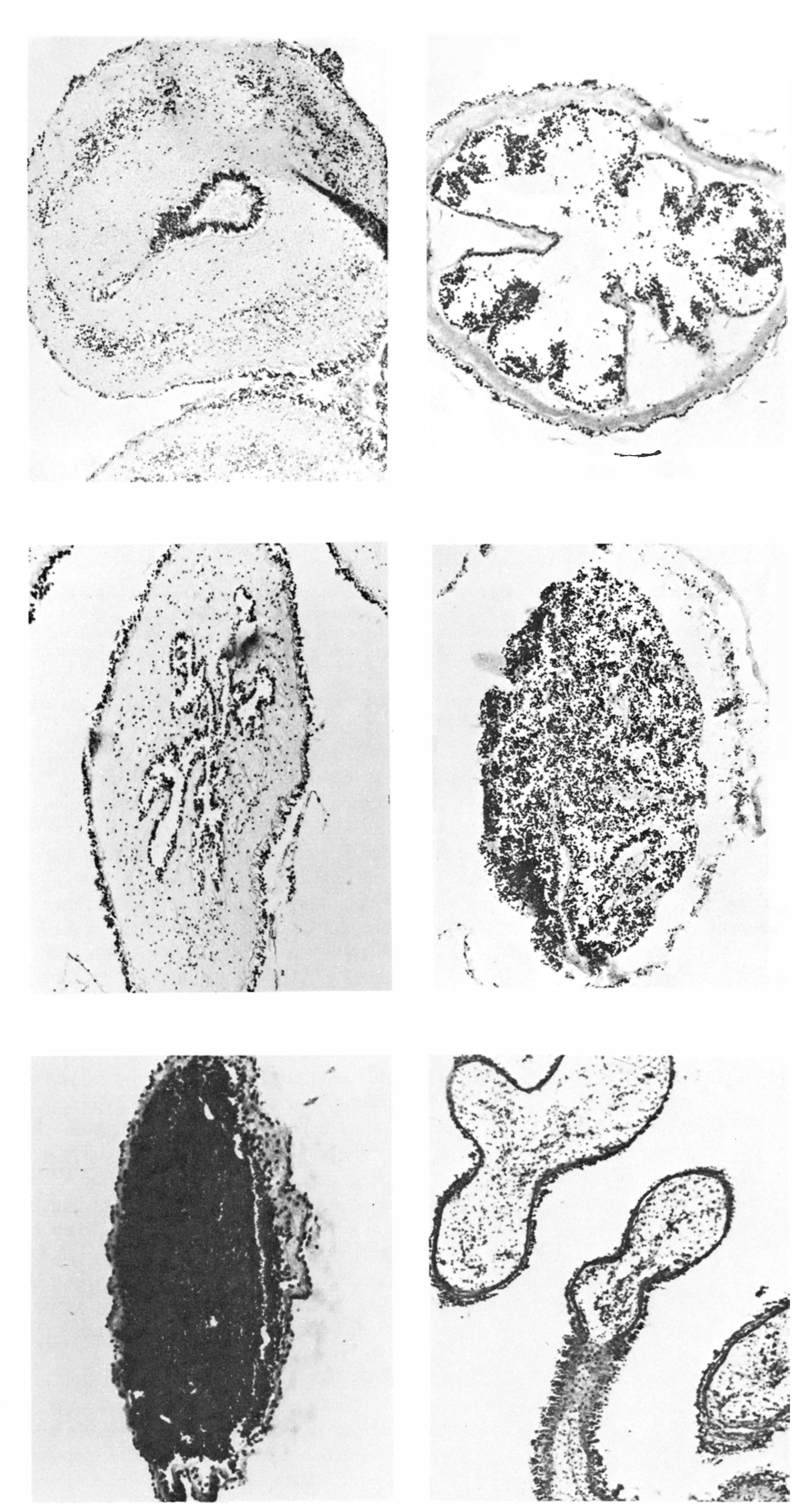
Nos hemos circunscrito hasta el momento al estudio bioquímico de la acción de la nucleasa micrococal sobre la cromatina. Veamos qué sucede cuando en experiencias similares se emplean otras nucleasas. El enzima DNasa II produce la misma serie regular de subunidades que la nucleasa estafilocócica al actuar en la cromatina, pero con una sola salvedad. La relación de multiplicidad de los fragmentos de ADN es distinta. Este enzima realiza incisiones en la fibra nucleínica a intervalos de 100 pares de bases nitrogenadas, a diferencia de la nucleasa micrococal que cortaba regularmente cada 200 parejas. Si se tiene en cuenta que la integridad del nucleosoma se mantiene y que los segmentos de 100 pares de bases sólo se detectan después de eliminar el complemento proteico mediante extracción y purificación del ADN de los productos de digestión, debe asumirse cierto grado de variación en la sensibilidad nucleásica del ADN. El punto medio del trozo de hebra de 200 pares de bases del ADN nucleosómico debe ser específicamente accesible a la DNasa II. Con todo, esta singular incisión no debe ser lo suficientemente profunda como para producir el desdoblamiento de la partícula, y sólo es reconocible al proceder a la extracción del ácido nucleico, obteniéndose entonces fragmentos de 100 pares de bases o múltiplos integrales de ellos.

Otro resultado sorprendente lo constituye la hidrólisis enzimática de la cromatina con DNasa I. Esta molécula en-

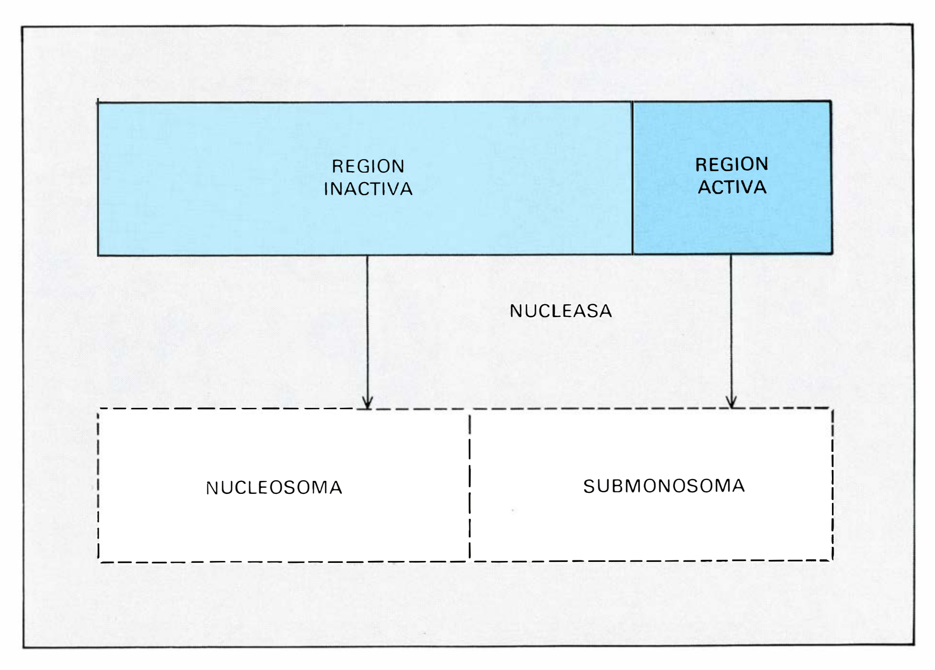
zimática secciona el ADN con una regularidad parecida a la de los otros dos enzimas mencionados, pero con una frecuencia mucho más elevada. El ADN liberado de los productos de la reacción con este enzima resulta en una serie regular de segmentos múltiplos de 10 pares de bases. La fibra de ADN cromatínico ofrece a este enzima puntos de acceso con una periodicidad de 10 parejas de desoxirribonucleótidos. Este dato es muy importante, ya que constituye el primer indicio acerca de la localización de los elementos integrantes del nucleosoma. El fácil acceso de este último enzima al ADN apoya de forma concluyente la posibilidad de que el ácido nucleico se sitúe en la parte exterior del nucleosoma. Parece así probable que la agrupación globular de histonas actúe de soporte de la doble hebra del ADN. La cadena de biopolímero se arrollaría así alrededor del octámero proteínico como en un ovillo. La colocación externa del ADN en el nucleosoma es mucho más idóneo para facilitar el normal desarrollo de sus funciones genéticas fundamentales. La elevada frecuencia de puntos de reconocimiento y acceso de la DNasa I lleva a desconfiar de un condicionamiento por parte de la secuencia del ADN. Esta debería poseer una regularidad desconocida para ofrecer señales a intervalos de 10 pares de nucleótidos. Parece mucho más coherente una motivación de tipo estructural. De nuevo el cálculo dimensional debido a las mediciones físicas nos señala que la doble hélice de ADN, al arrollarse, invierte alrededor de 10 pares de bases por vuelta. En estas condiciones no es de extrañar que los puntos más sensibles de la cadena fuesen los extremos de cada enrollamiento distanciados precisamente de 10 nucleótidos.

En resumen, los enzimas nucleolíticos actúan sobre la cromatina produciendo una serie de partículas compuestas por una subunidad repetitiva fundamental, el nucleosoma. Ahora bien, la periodicidad del acceso enzimático varía según el tipo de nucleasa. Las que ostentan mayor frecuencia de reconocimiento ponen en evidencia peculiaridades estructurales intranucleosómicas demostrativas de la presencia de regiones específicas, de mayor sensibilidad, a lo largo del segmento de ADN.

Señalábamos antes que, en condiciones enérgicas de hidrólisis con nucleasa micrococal, el fragmento de ADN de 200 pares de bases se reducía a 140. A título de explicación indicábamos que ello era probablemente debido a una degradación progresiva provocada por el enzima



ANALISIS HISTOLOGICO de gónadas masculinas de *Holothuria* en diversas fases de la maduración germinal; muestra cómo en el momento de la activación, el epitelio germinal aparece como una capa circular en el centro del túbulo. Este estrato invade progresivamente la cavidad tubular hasta rellenarlo completamente de espermatozoides en la culminación de la espermatogénesis. Efectuado el desove, el tejido conectivo recubre nuevamente el interior de los túbulos iniciándose la recesión a la fase de inactividad germinal. Las distintas fotografías se han tomado a 150 aumentos.



LA EXISTENCIA de partículas subnucleosómicas en la cromatina plantea el problema de si sus regiones activas e inactivas poseen idéntica organización estructural. Las pruebas de que se dispone tienden a sugerir que la superestructura en nucleosomas corresponde a las zonas genéticamente inactivas mientras que las partículas submonosómica proceden de cromatina transcriptora.

que en un momento determinado comenzaba a digerir el propio ADN nucleosómico. Diversos laboratorios han examinado con detalle los productos de ADN obtenidos por digestión con nucleasa de cromatina de diversos tejidos celulares. Estos estudios han revelado que el fragmento de ADN nucleosómico de 200 pares de bases no es constante, sino que existe un margen de variación dependiente de la especie celular analizada. En contraposición, el tamaño del ADN correspondiente a los productos finales de las digestiones es sistemáticamente el mismo y equivale a un segmento de cadena nucleotídica de 140 pares de bases nitrogenadas.

Estos resultados hacen suponer que el nucleosoma posee dos partes diferenciadas. Una de ellas correspondería al núcleo de la partícula constituido por el oligómero de histonas y la hebra de ADN de 140 pares de bases en íntima asociación. El resto del ADN pertenecería al espaciado entre dos núcleos monoméricos adyacentes. Una tal disposición es coherente con los datos experimentales. Al comenzar el ataque enzimático se producen nucleosomas que contienen unos trozos de ADN sueltos a modo de apéndices. Estos fragmentos, al no interaccionar directamente con las histonas del núcleo, son más sensibles al enzima y, por tanto, son recortados con el progreso de su acción. Finalmente, resulta el núcleo particulado conteniendo el segmento de 140 pares de bases y sus histonas asociadas.

La relación ADN/histonas en todos los organismos tiene un valor aproximado que se acerca a la unidad. Con todo existen ciertas desviaciones según las especies debido a la diferencia del contenido. Esta circunstancia explicaría que algunas células, al poseer una cantidad ligeramente menor de histonas, tuviesen una longitud de ADN libre superior y, por tanto, diesen lugar a monómeros dotados de un segmento mayor de ADN. A estas variaciones de contenido de histonas y ADN es atribuible la existencia de diferencias en la distancia repetitiva del ADN nucleosómico. En cambio, la constancia del núcleo monomérico viene impuesta por la estricta interacción en el interior del nucleosoma que involucra a la agrupación globular de histonas y al correspondiente fragmento de ácido nucleico invariable y fijo. Este supuesto, en consecuencia, nos lleva a distinguir dos tramos específicos de ADN en el nucleosoma. De una parte, la hebra estrictamente intranucleosómica y, por la otra, la cadena de interconexión o espaciadora de nucleosomas adyacentes.

Hasta aquí hemos hecho hincapié en consideraciones de índole meramente estructural, pero es muy probable que exista una estrecha correlación entre las variaciones descritas en el espaciado internucleosómico y la actividad genética. En su momento se ha omitido indicar que existe una graduación precisa entre las diferencias de longitud observadas del ADN espaciador. Cuanto más primitiva es la célula evolutivamente –por ejemplo los eucariotas inferiores, como la levadura– menor es la longitud del ADN interconectador. De idéntica forma sucede respecto del grado de actividad de la especie celular. Las células de tipo somático exhiben un espaciado nucleosómico mayor que las células genéticamente activas. Parece, pues, real la correlación entre el tamaño del ADN internucleosómico y la actividad genética. Cabe esperar que esta periodicidad sea más compleja en cromatinas pertenecientes a células activas, con regiones transcriptoras específicamente diferenciadas, que en cromatinas inactivas, donde el objetivo único consiste en un eficaz empaquetamiento del ADN para su conservación. No hay que olvidar la posible intervención de la histona H1 en este proceso. Esta proteína, que presenta cambios durante el desarrollo y diferenciación celulares, puede de igual manera contribuir a la modificación de la accesibilidad enzimática a la cromatina activa. En definitiva, hemos avanzado sensiblemente en el conocimiento de la organización estructural de la cromatina, habiendo establecido la naturaleza y gran parte de las características del nucleosoma juntamente con la distribución periódica a lo largo de las fibras, pero las funciones asignadas a esta subunidad fundamental permanecen en gran parte todavía por dilucidar. El problema de si la estructura particulada de la cromatina que acabamos de describir se mantiene en las regiones activas en transcripción no está resuelto completamente. Las pruebas de que disponemos en la actualidad tienden a sugerir que no es así.

El régimen de actividad genética adquiere singular intensidad y relieve durante el fenómeno de la diferenciación celular. En consonancia con las distintas y complejas fases de este proceso biológico, la cromatina experimenta una serie de profundas alteraciones cuya significación constituye un objeto de examen particularmente atractivo y de gran interés. Llevado por la trascendencia del tema, el equipo de investigación de nuestro laboratorio últimamente ha dirigido su atención hacia él a través de diversas líneas de trabajo, sobresaliendo con especial énfasis el estudio bioquímico acerca de la repercusión y dependencia del desarrollo celular en el ordenamiento estructural del cromosoma. Parte de este estudio que describiremos someramente se ha realizado conjuntamente con E. Rocha sobre la acción de la nucleasa micrococal en cromatina de células germinales masculinas en desarrollo del equinodermo *Holothuria tubulosa*. La razón del empleo de un tejido germinal radica en la implicación fundamental de las

transiciones de la cromatina a lo largo del proceso de maduración de las células. La selección de este organismo fue debida a que, en contraste con gran variedad de organismos superiores, no experimenta cambios espectaculares en la dotación de proteínas nucleares básicas durante la espermiogénesis, es decir, esta especie animal exhibe en su núcleo un espectro de transiciones limitado. La condensación y contracción de la cromatina que tienen lugar en la maduración del esperma no son acompañadas del desplazamiento de las histonas por proteínas más básicas del tipo de las protaminas. En consecuencia, las variaciones observables en la respuesta de la cromatina a la nucleasa a través del desarrollo deben involucrar como mínimo una parcial modificación del ordenamiento estructural del complejo de desoxirribonucleohistona, o sea, del nucleosoma.

Para una mejor comprensión del comportamiento de la cromatina durante la maduración germinal es aconsejable una esquemática descripción del sistema reproductor de esta especie. Cada individuo posee una sola gónada localizada en la parte anterior de la cavidad celómica y constituida por numerosos túbulos unidos en forma de racimo por la base. En la gónada inmadura dichos túbulos son usualmente cortos, pero experimentan un drástico incremento de tamaño durante el período de maduración sexual, que se refleja claramente en el incremento desmesurado del propio peso de la gónada. El conjunto de las variaciones testiculares indica que el ciclo reproductivo de la *Holothuria* exhibe una cadencia anual con una sola fase de desove por ciclo. El análisis histológico nos muestra que los túbulos gonadales están recubiertos externamente de epitelio celómico que rodea una capa de fibras musculares, tanto circulares como longitudinales. En la gónada en reposo la superficie interior está constituida por tejido conjuntivo. En el momento de la activación, el epitelio germinal aparece en la luz del túbulo como un pequeño estrato circular que se agranda progresivamente con la maduración y desarrolla lobulaciones e invaginaciones a medida que se extiende hacia el exterior, desplazando al tejido conjuntivo. Los diversos tipos de células germinales que intervienen en la espermatogénesis se originan en esta capa, distribuyéndose los tipos original e intermedios en la parte externa mientras que los espermatozoides se localizan en el interior. Una sobresaliente característica puesta en evidencia por los exámenes histológicos es la persistente abundancia de espermatocitos a lo largo de todo el ciclo, incluso en fases avanzadas de la maduración. La duración del período de crecimiento y la tardía aparición de esperma maduro en las gónadas son, probablemente, una consecuencia directa de la reiterada presencia de este tipo celular en la mayoría de las etapas del desarrollo. En síntesis, las peculiaridades del ciclo reproductivo en esta especie animal sugieren la existencia de una íntima conexión entre el régimen de actividad genética y las modificadiones de la cromatina a lo largo del desarrollo.

Veamos ahora cómo las alteraciones morfológicas se reflejan a nivel molecular de la cromatina de *Holothuria.* En primer lugar, el acceso de la nucleasa micrococal a la cromatina de gónadas inmaduras resulta afectado por las condiciones de preparación. El núcleo celular, que debería ser normalmente muy permeable al enzima, ofrece en este organismo una resistencia inusitada a la penetración enzimática. Este comportamiento se acentúa con la presencia de iones calcio en el medio de aislamiento nuclear añadidos con objeto de minimizar su fragmentación, a pesar de ser este catión indispensable para la acción de la nucleasa. Es necesario liberar la cromatina del núcleo para que el enzima pueda iniciar la solubilización del ADN aunque con marcada lentitud. La única alternativa eficaz para facilitar el proceso hidrolítico consiste en someter la cromatina a fuerzas de cizallamiento. Otra relevante observación efectuada por E. Rocha durante la experimentación en el laboratorio es que el régimen de digestión enzimática decrece a medida que avanza el ciclo de maduración alcanzándose valores mínimos en las últimas etapas espermatogénicas. Este comportamiento decididamente anómalo de la cromatina nuclear frente a la nucleasa micrococal implica una singular compacidad en la organización de la membrana nuclear que reduzca su permeabilidad y entorpezca el acceso del enzima. Esta conducta sugiere además que el complejo de nucleohistona en las células germinales posee un grado de cohesión extremo que se traduce en una marcada ocultación del ADN.

El análisis de sedimentación de digeridos enzimáticos revela la existencia de nucleosomas, juntamente con agrupaciones de tamaño mayor. Con todo, el hecho notable es la presencia de un material ligero de lenta sedimentación cuya proporción depende de la concentración enzimática. En condiciones de digestión restringentes, las especies multiméricas se desdoblan, como era previsible, en nucleosomas, descartándose la posibilidad de que el origen del material subnucleosómico mencionado se debiera a degradación del nucleosoma. Concomitante con la disminución descrita del grado de solubilización en función del desarrollo hemos comprobado la progresiva desaparición de la población de subnucleosomas hasta su total ausencia de la cromatina de células espermáticas maduras.

La conversión de las formas multiméricas en nucleosomas es coherente con la relación precursor-producto postulada. Ahora bien, la variable accesibilidad de la nucleasa a la cromatina detectada en el tejido germinal de esta especie durante el ciclo reproductivo, así como la presencia inicial de subnucleosomas y su posterior desaparición en las fases últimas de la maduración, en las que el grado de condensación de la cromatina es uniforme, son demostrativas de la generación de modificaciones estructurales, en el sentido de que las regiones de cromatina en las que se originan los subnucleosomas sufren una alteración en un momento dado que las convierte en inasequibles al enzima. La coexistencia de regiones cromatínicas con un ordenamiento estructural distinto, puesto de relieve por la acción diferencial de la nucleasa, carece de sentido sino es reflejo de estados funcionales dispares.

En resumen, nuestra observación de que nucleosomas y sus agrupaciones multiméricas son los productos exclusivos de la digestión por nucleasa micrococal al término de la espermatogénesis en la *Holothuria,* cuando su cromatina aparece condensada por completo e inactiva totalmente, arguye en favor de que el modelo propuesto de organización del ADN cromosómico en su nivel molecular más íntimo, basado en una distribución periódica integrada por una subestructura repetitiva fundamental, el nucleosoma, uniformemente repartida a lo largo de las fibras de cromatina, se halle restringido a formas genéticamente inactivas.

Es evidente que muchos aspectos de la naturaleza molecular del cromosoma restan todavía confusos, pero la correlación entre el ordenamiento estructural y la función genética empieza a revelarse. La especulación acerca de los mecanismos de mantenimiento y control del estado del cromosoma y su influencia en los procesos regulatorios de la expresión genética va cediendo ante los hallazgos experimentales. Sin duda alguna, al término de este constante progresar de la biología molecular se halla el conocimiento profundo de la esencia de los fenómenos que constiuyen la vida.

El sistema de apareamiento de lek del gallo de las artemisas

En un sistema de lek un gran porcentaje de hembras se aparea con un reducido porcentaje de machos. Se examina con detalle el modo en que actúa tal sistema en el gallo de las artemisas de las montañas Rocosas

R. Haven Wiley, Jr.

De todos los rituales extravagantes en el cortejo animal, pocos son tan impresionantes como el sistema de apareamiento de lek del gallo de la artemisas *(Centrocercus urophasianus)*. En las elevadas mesetas cubiertas de artemisas de las montañas Rocosas ambos sexos se congregan durante la estación de reproducción en terrenos de exhibición comunales, llamados leks (originalmente una palabra escandinava que significa juego). Allí, los machos se exhiben reiteradamente. Aunque en un lek pueden reunirse 50 o 60 machos de gallo de las artemisas, la mayoría de las hembras visitantes copulan sólo con uno o dos machos, siempre los mismos; el gran número restante de éstos no encuentran compañera.

El sistema de apareamiento del lek, caracterizado por una asociación transitoria de ambos sexos en un lugar regular y por un éxito copulador extremadamente desigual entre los machos, se ha desarrollado entre especies de aves emparentadas muy lejanamente, entre las cuales se cuentan ciertos tetraónidos (urogallos y lagópodos), aves del paraíso, correlimos y andarríos, tejedores y manaquines y cotingas de América Central y del Sur. Se encuentra también en algunos insectos, peces y mamíferos. Por ser este sistema un caso tan claro de apareamientos desiguales por parte de los machos ha ejercido desde siempre cierta fascinación en los zoólogos. Pensé que un ejemplo tan extremo como el del gallo de las artemisas podría clarificar algunos principios generales de los sistemas de apareamiento de los animales, y por esta razón emprendí recientemente un estudio de tres años de duración sobre el apareamiento del gallo de las artemisas.

En invierno, cuando las mesetas de artemisas se hallan azotadas por la nieve y las ventiscas, este grupo de gallos dependen casi por entero de una planta perenne, la artemisa mayor, tanto para refugiarse como para comer. Dichos tetraónidos se cuentan entre las mayores aves de su familia y presentan un dimorfismo sexual acusado: los machos adultos pesan una media de 2,5 kilogramos, y las hembras adultas sólo 1,2 kilogramos. En los primeros días cálidos de febrero los machos empiezan a congregarse en leks dispersos en las llanuras de artemisa. De marzo a abril y hasta entrado mayo, cada macho se dirige, una mañana tras otra, a su lek particular y, dentro de éste, a una posición determinada. Las hembras visitan el lek durante un período mucho más breve, por lo general unos 20 días en abril.

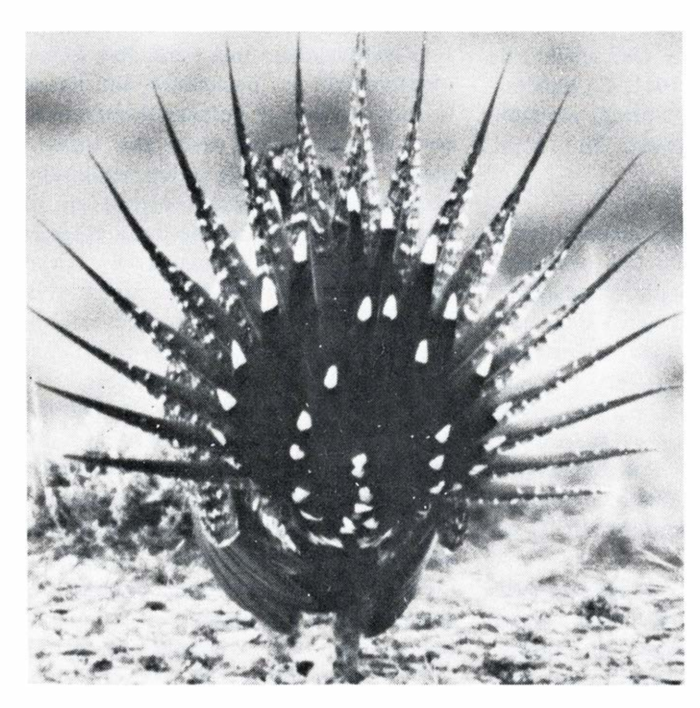
La actividad del lek empieza con el alba, aproximadamente una hora antes de la salida del sol, y continúa a lo largo de tres o cuatro horas. Durante este período, los machos pasan mucho tiempo repitiendo lo que los primeros observadores de los gallos de las artemisas llamaron "pavoneo". El macho hincha su saco esofágico elástico elevándolo y dejándolo caer después. Repitiendo esta operación dos veces, expande el saco hasta que éste contiene unos cuatro litros de aire. Luego, contrayendo los músculos superficiales del pecho, comprime el saco hinchado y expulsa de golpe el aire comprimido. El resultado es un sonido parecido a un chasquido resonante, muy similar al que produce un tapón al descorchar una botella. Cada ciclo de hinchamiento y disparo dura poco más de tres segundos; entre pavoneos sucesivos, el macho suele colocarse en una actitud conspicua, con las blancas plumas del cuello erizadas y con las plumas caudales erguidas verticalmente. Las exhibiciones complicadas son típicas de los animales que utilizan leks para aparearse, pero el gallo de las artemisas llega al paroxismo del comportamiento grotesco.

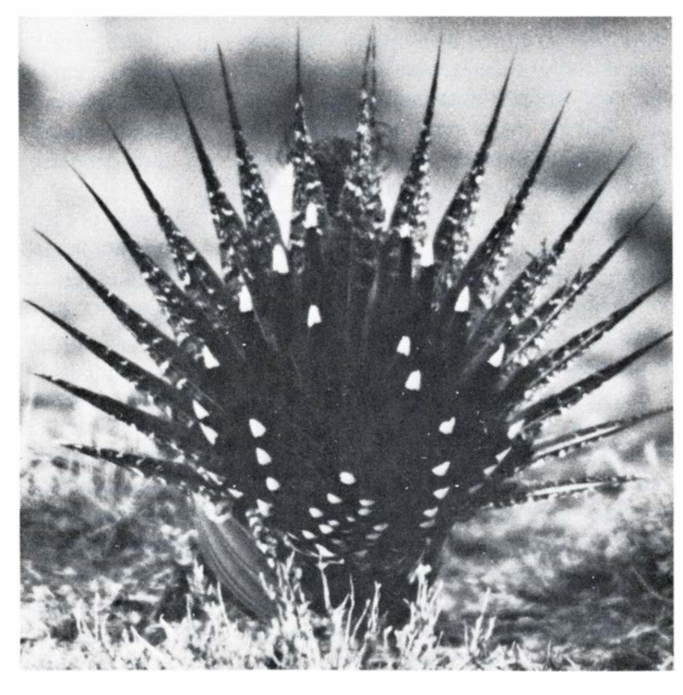
Cuando las gallinas de las artemisas visitan el lek en abril, llegan en gran número, hacia la salida del sol, y tienden a congregarse en una banda densa en el centro. El único macho, o los dos machos que se encuentran cerca del centro se pavonean adelante y atrás en medio de la congregación de hembras y copulan a intervalos con ellas. Cada hembra suele copular sólo una vez y después se marcha para hacer el nido, a veces a una distancia de hasta cuatro kilómetros del lek. La hembra pone de seis a ocho huevos, los incuba durante 26 días por término medio y cuida de los pollos sin ninguna asociación posterior con ningún macho.

De los 50 o 60 machos que concurren en un lek durante la estación de reproducción, la inmensa mayoría nunca copula en absoluto, y sólo unos pocos se aparean repetidamente. A lo largo de una estación un macho afortunado monta de 20 a 60 hembras; una vez observé 34 cópulas fructíferas realizadas por el mismo macho en una sola mañana. En conjunto, el 90 por ciento, al menos, de las cópulas realizadas en un lek determinado lo son por no más de un 10 por ciento de los machos presentes.

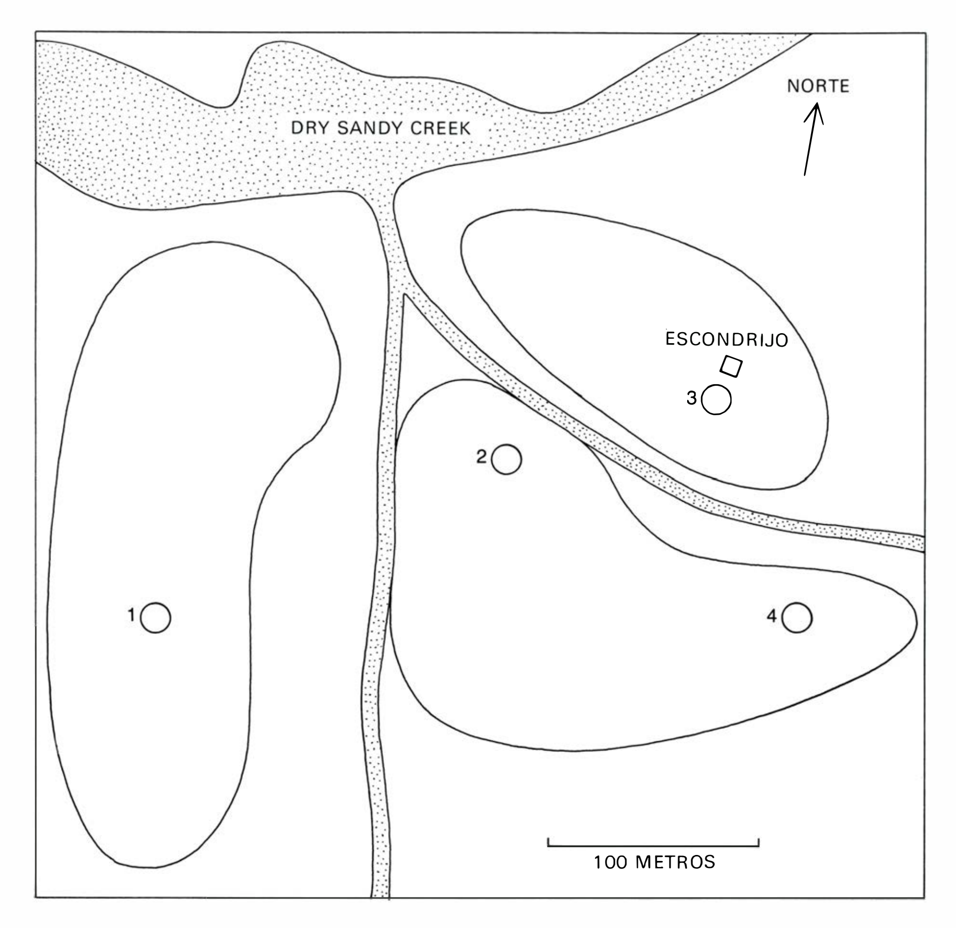
Los antiguos naturalistas que se interesaban por la teoría darwiniana de la selección sexual se encontraban particularmente intrigados por el sistema de apareamiento de lek. Por ejemplo, cuando Edmund Selous describió los leks del gallo lira en Inglaterra, en 1909, destacó la competencia entre machos y la selección que las hembras visitantes hacían de los individuos más vigorosos. De hecho, y durante las décadas que transcurrieron

EL CENTRO DE COPULA en un terreno de exhibición comunal, o lek, de los gallos de las artemisas, en Wyoming (Estados Unidos), está abarrotado de hembras. Un gallo macho, mucho mayor que las hembras, parece en esta fotografía incluso mayor de lo normal, pues ha hinchado su saco esofágico con aire en el curso de una exhibición de cortejo llamada pavoneo. Esconde la cabeza en una gorguera de plumas collares blancas.





ABANICO ERGUIDO de las plumas caudales. Al levantar las plumas caudales, las plumas pardas, más cortas y situadas bajo aquéllas, se hacen visibles; sus puntas blancas forman un modelo característico que difiere de un macho a otro, como se aprecia en estas fotografías. Recordando los distintos modelos el autor pudo identificar a cada uno de los machos que divisaba desde su escondrijo y seguir sus movimientos.



UN GRAN LEK EN WYOMING comprendía tres zonas en las que los machos ocupaban territorios; estas zonas tenían en total cuatro centros de cópula, cada uno de ellos de unos 50 metros cuadrados de extensión. La mayoría de las hembras se congregaban en estos centros. Por las mañanas, y a su debido tiempo, más de 250 machos se reunían en el lek. El escondrijo del autor se instaló cerca de uno de los cuatro centros de apareamiento *(arriba, a la derecha)*. El Dry Sandy Creek, de donde ha tomado su nombre el lek, es alimentado por arroyos tributarios.

antes de que se dispusiera de datos sobre el comportamiento poligínico en otros animales, las aves que formaban leks proporcionaron los principales ejemplos para la teoría de la selección sexual. Hoy día, con estudios detallados, estamos en una posición mejor para comprender las interacciones sociales que genera la desigual distribución de los apareamientos en los leks. Mis propias investigaciones ponen de manifiesto que el gallo de las artemisas presenta muchos rasgos de comportamiento que se encuentran en otros animales que forman leks, aunque de un modo menos exagerado.

Característica de los leks es su localización tradicional. Los gallos de las artemisas retornan uno y otro año virtualmente al mismo punto. En 1949, Robert Patterson, de la Wyoming Game and Fish Commission, localizó todos los leks de gallos de las artemisas en 650 hectáreas de artemisa en la región occidental del estado. Repitió su estudio en 1950 y 1951. No halló nuevos leks, pero observó que éstos se encontraban cada año en la misma situación. El récord documentado de querencia a una localidad determinada corresponde al lek de Muddy Springs, cerca de Laramie (Wyoming), estudiado por primera vez por John Scott, de la Universidad de Wyoming, en 1940 y 1941. La localización del lek de Muddy Springs no había cambiado cuando yo trabajé en la zona, 28 años después.

No sólo el propio lek se encuentra en el mismo lugar un año tras otro, sino también el centro de cópula, una pequeña área interior de unos 10 metros de diámetro, en la que tienen lugar la mayoría de las cópulas. Scott, Patterson y yo mismo hemos observado esta persistencia en la localización, si bien la situación del centro de cópula parece perder esa persistencia en leks pequeños o en aquellos que se han visto perturbados por las actividades humanas.

Los leks de los gallos de las artemisas varían en cuanto al número de machos que los ocupan. En algunos leks pequeños sólo se congregan unos cuantos machos; en otros mayores hay más de 400 machos, aunque un número tan elevado es insólito. En cualquier caso, cuando se congregan más de 50 o 60 machos, el lek siempre tiene más de un centro de cópula. En un lek realmente grande puede haber un centro de cópula por cada 50 machos aproximadamente. Esto significa de hecho que un lek mayor es un agregado de otros más pequeños, de modo que la unidad básica de la organización social del gallo de las artemisas es un lek de unas dos hectáreas de superficie que es visitado a su tiempo por hasta 50 o 60 machos y tiene un centro de cópula único de unos 50 metros cuadrados de extensión.

Incluso para un biólogo que tirita en su escondite al amanecer, solo en las vastas llanuras de Wyoming, un lek grande es un espectáculo impresionante. Cada gallo macho que llega se instala en su territorio en el interior del lek, una mañana tras otra, para pavonearse y defender sus límites contra la intrusión de otros. Si ocurre que un macho se pavonea cerca del límite del territorio de otro, el macho desafiado suele responder desplazando su propio pavoneo más cerca del límite amenazado. Periódicamente los machos vecinos interrumpen su pavoneo y lo sustituyen por un combate directo en la frontera. Estos combates pueden empezar de súbito, al arremeter un macho contra otro que se ha acercado demasiado a la divisoria. Los dos contendientes acaban situándose uno casi al lado del otro, costado contra costado, pero con dirección frontal opuesta; al mismo tiempo, se desplazan hacia delante y hacia atrás unos cuantos pasos cada vez.

Puede ocurrir que la tensión aumente y se desencadene una verdadera lucha. Manteniendo sus posiciones inmediatas, los dos machos se golpean mutuamente con vigorosos golpes de las alas. Lo normal es que ninguno de ambos resulte claramente vencido; como ocurre en los combates territoriales de otras especies animales, los antagonistas se retiran pronto a sus dominios respectivos. Aunque pueden llegar a herir al adversario, nunca vi a un macho dañado como consecuencia de uno de estos combates.

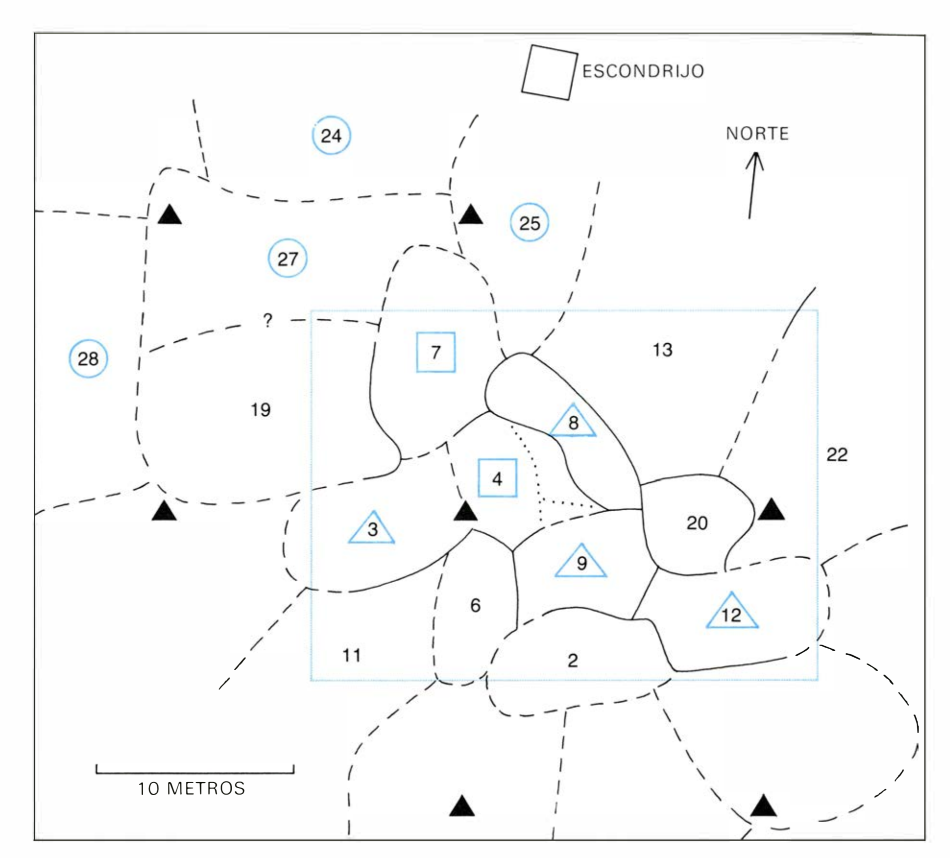
La observación detallada de este tipo de comportamiento precisa de la habilidad para identificar individualmente a cada macho. En vez de arriesgarme a asustar a las aves capturándolas para marcarlas artificialmente, aprendí a identificar a los individuos mediante el modelo único que forman las plumas situadas bajo su cola erguida; las cortas plumas subyacentes a la cola son de color pardo chocolate con las puntas blancas, y no existen dos modelos que sean exactamente los mismos. En los años que

duró mi estudio, cada temporada me resultó fácil reconocer a los distintos machos que divisaba desde mi escondite.

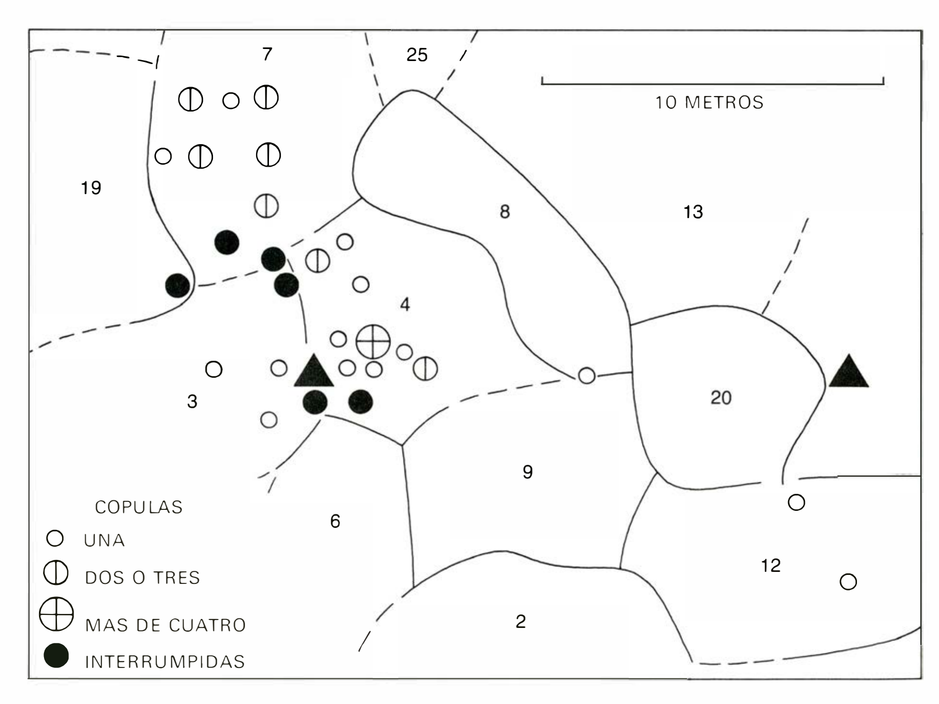
También pude trazar la localización de cada combate fronterizo en un radio de 30 a 50 metros de mi escondrijo, levantando un mapa detallado del terreno; este mapa incluía cada conjunto de artemisa y cada grupo de gramíneas dentro del campo visual. El trazado del mapa resultó más fácil colocando una red de pequeñas banderolas cerca de mi puesto de observación. Me serví de una cámara filmadora de 16 milímetros para registrar las actividades de los machos. Filmando a un ritmo de una imagen cada cuatro segundos, pude registrar el comportamiento de hasta cinco machos durante un período de dos horas desde la salida del sol. Cuando dibujé sobre el mapa la situación de los combates fronterizos, resultó que las líneas de demarcación entre los territorios de machos adyacentes eran precisas: muchas fronteras estaban definidas con un metro o menos de variación. Los territorios situados cerca del centro de cópula eran menores que los situados en la periferia del lek: generalmente no superaban los 30 metros cuadrados, en comparación con los 100 metros cuadrados de los mayores.

De acuerdo con mis observaciones, cada macho de gallo de las artemisas es soberano en el interior de su propio territorio. Más importante todavía, en el interior de su territorio un macho puede copular sin verse desafiado por otro vecino. Sólo cuando un macho monta a la hembra cerca del límite territorial, el macho vecino se precipita sobre él y se entabla una batalla que interrumpe el apareamiento. El macho, independientemente de su éxito o de su fracaso en el apareamiento con las hembras, se halla sujeto al ataque del vecino si intenta acoplarse demasiado cerca del límite territorial de éste.

Ocupar un territorio en el interior o cerca del centro de cópula de un lek representa gozar de una gran ventaja en el apareamiento; un macho con esta localización copulará repetidamente, mientras que sus vecinos inmediatos lo harán sólo en contadas ocasiones. A pesar de esta discrepancia, el vecino menos favorecido no se dirigirá al macho más afortunado para interrumpir la cópula, y éste no se introducirá en el territorio del menos favorecido, a menos que uno u otro intenten el apareamiento demasiado cerca del límite territorial. Resulta evidente que un gallo macho tiene éxito en su apareamiento no debido a su ca-



LOS TERRITORIOS DE LOS MACHOS que rodean uno de los centros de cópula en el lek Dry Sandy estaban ocupados por unos 20 gallos macho en la época en que el autor hizo sus observaciones. A cada ave se le asignó un número de identificación; los cuatro números encerrados en los círculos de color corresponden a machos de un año en su primera temporada en un lek. Los dos números encerrados en los cuadrados de color identifican a los machos sexualmente más activos; los cuatro números de los triángulos de color corresponden a machos que también copulaban, pero lo hacían con menor frecuencia. Los triángulos negros corresponden a la localización de banderolas situadas siguiendo un cuadriculado y que ayudaban a la cartografía de los territorios; las fronteras recortadas con líneas de trazos discretos son las que se conocen de manera más imperfecta que las que se señalan en línea continua. El rectángulo de color delimita el área del lek que se muestra ampliada en la ilustración inferior. Unicamente los seis machos así identificados consiguieron aparearse, mientras que no copularon las otras 14 aves señaladas en la figura, ni ninguno de los numerosos machos poseedores de territorio más alejados del centro.



EL REGISTRO DE APAREAMIENTOS a lo largo de un período de 20 días en abril muestra una gran actividad en los machos números 7 y 4, una actividad menor en los machos 3, 9 y 12 y ninguna actividad copuladora en el número 8. Se interrumpieron seis intentos de apareamiento cuando el macho pretendió aparearse en un punto demasiado cercano a la frontera de un territorio vecino.

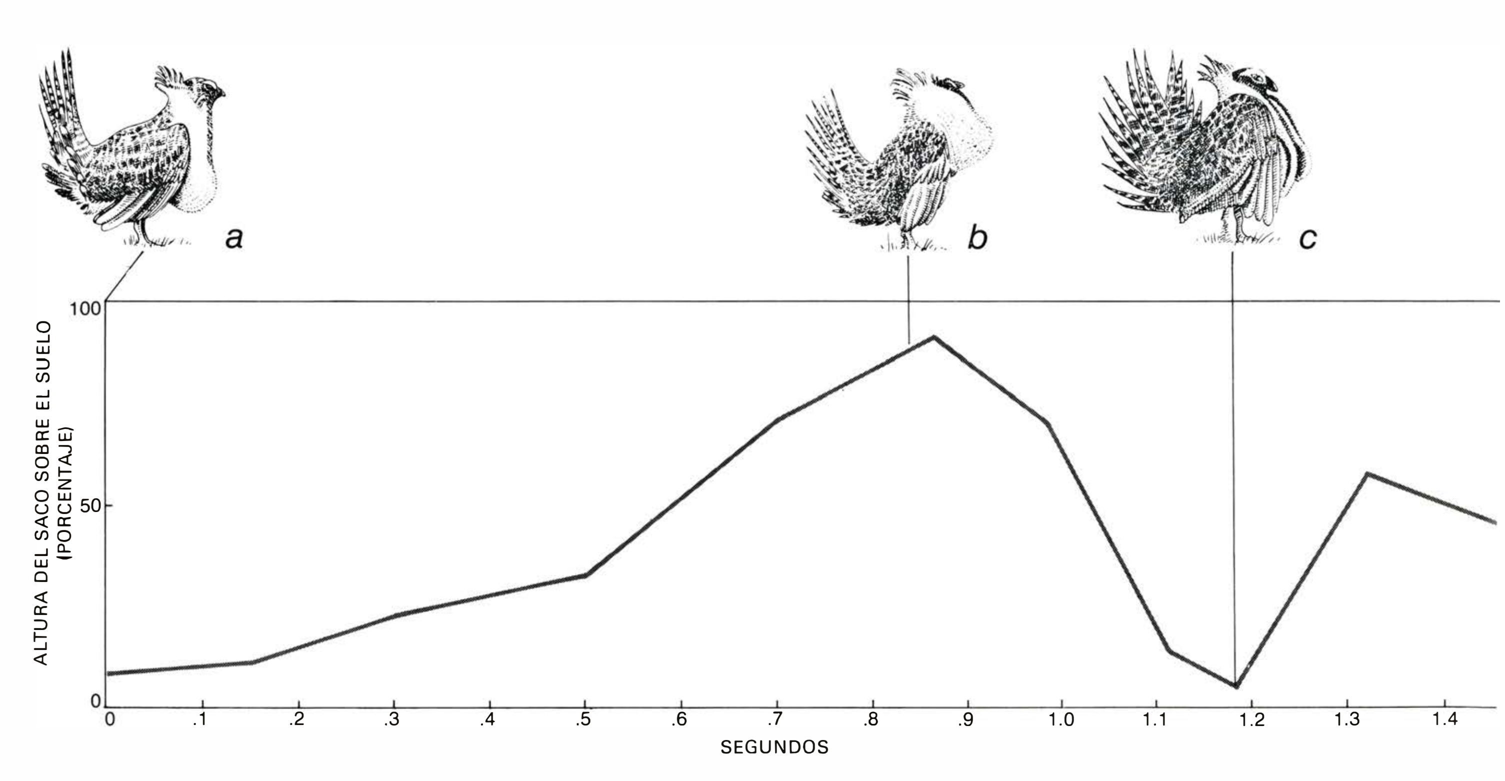
pacidad de interrumpir los intentos copulatorios de otros machos, sino porque ha obtenido un territorio que coincide con un centro de cópula.

Antaño los naturalistas no sabían exactamente si los machos de los gallos de las artemisas demostraban un comportamiento territorial en los leks, comportamiento en el que los vecinos se enfrentan como iguales en fronteras mutuas, o bien presentaban una jerarquía de dominancia, en la que los machos dominantes niegan a los subordinados la oportunidad de aparearse. De hecho, las interacciones sociales de los gallos de las artemisas son una mezcla de estas dos pautas clásicas de comportamiento. Los machos vecinos se enfrentan en la línea de demarcación más o menos como iguales. Sólo ocasionalmente se infiltran en el territorio del otro, aunque un macho se apare frecuentemente y los demás no. Con todo, los machos son atraídos al centro de cópula del interior del lek. De ello resulta que cada macho tiende a pavonearse durante la mayor parte del tiempo cerca del límite de su territorio más cercano al centro de cópula. Al mismo tiempo evita las intrusiones de los vecinos cuyos territorios se hallan más alejados de dicho centro que el suyo. Comprobé que cada macho ataca a los vecinos más alejados del centro de cópula con más frecuencia que el caso inverso, es decir, que los vecinos ataquen al macho más central. Efectivamente, los machos practican un "territorialismo polarizado": su rango en el seno de una jerarquía de dominancia se establece según la distancia de su territorio al centro de cópula.

¿Qué respuestas de comportamiento por parte de las hembras establecen el centro de cópula? Una hipótesis, propuesta por los primeros naturalistas que estudiaron los leks, es que los machos difieren entre sí y que las hembras, según determinado criterio de forma o de comportamiento, eligen para aparearse al macho más atractivo del lek, y, por tanto, transforman su territorio en centro de cópula. A primera vista, la hipótesis parece plausible: el observador casual de un lek pronto advierte que los machos adyacentes al centro de reproducción se pavonean con más frecuencia que los machos más alejados, dedican más tiempo a la exhibición y tienen asimismo combates fronterizos más frecuentes con los machos vecinos. Es concebible que una de estas acciones, o el conjunto de las mismas, pueda atraer a las hembras. Sin embargo, hay que mostrarse precavido a la hora de aceptar esta explicación. No puede excluirse la posibilidad de que la mayor actividad de los machos más cercanos al centro sea el resultado, y no la causa, de congregarse allí las hembras. Si los machos se muestran más activos cuando están cerca de las hembras, entonces, y sin importar el modo en que las hembras seleccionan un lugar para congregarse, los machos del sitio escogido exhibirán una mayor actividad.

Para probar estas explicaciones alternativas registré la actividad de los machos situados en posiciones centrales o periféricas del lek en condiciones de proximidad a las hembras equivalentes. Las secuencias de fotografías me permitieron documentar el comportamiento de los machos en tres circunstancias distintas: cuando las hembras se encontraban dentro de los límites territoriales del macho en observación, cuando las hembras estaban dentro del territorio de un macho vecino adyacente y cuando no había hembras en el territorio del macho observado ni en el de ninguno de sus vecinos inmediatos.

En el primer caso observé que los machos se pavoneaban con frecuencia, a menudo a un ritmo aparentemente máximo (más de seis pavoneos por minuto). Se pavoneaban con menor frecuencia cuando las hembras sólo se encontraban en el territorio de un vecino, y casi nunca lo hacían si no había hembras en los territorios adyacentes. De modo que no había diferencias notables entre el comportamiento de los machos con éxito y el de los machos sin él en condi-



SECUENCIA DEL PAVONEO del macho de gallo de las artemisas durante una sola exhibición, de tres segundos de duración. El gráfico inferior indica las variaciones en la altura del saco esofágico del macho, expresada como porcentaje de la máxima altura; en la figura *d* se obser-

ciones de proximidad equivalente a las hembras.

Jon Hartzler, de la Universidad de Montana, ha intentado asimismo descubrir diferencias individuales entre machos de gallos de las artemisas, con éxito y sin él, a distintas distancias de las hembras. Empleó el criterio de la distancia absoluta entre el macho y la hembra en vez de apoyarse en la presencia, o ausencia, de la hembra en el territorio del macho. Encontró que los machos con éxito son algo más activos que los que no lo tienen, a una misma distancia de las hembras. La diferencia en nuestros hallazgos respectivos puede deberse a los distintos criterios seguidos en lo que respecta a la proximidad de un macho a las hembras. Por el momento, no está claro si las diferencias de comportamiento de los distintos machos afectan a la elección que las hembras hacen de un lugar para aparearse. Sigue permaneciendo el hecho de que la actividad de pavoneo más frecuente de los gallos de las artemisas adyacentes al centro de cópula del lek es, ante todo, el efecto de que las hembras se congregan aquí, en lugar de ser la causa.

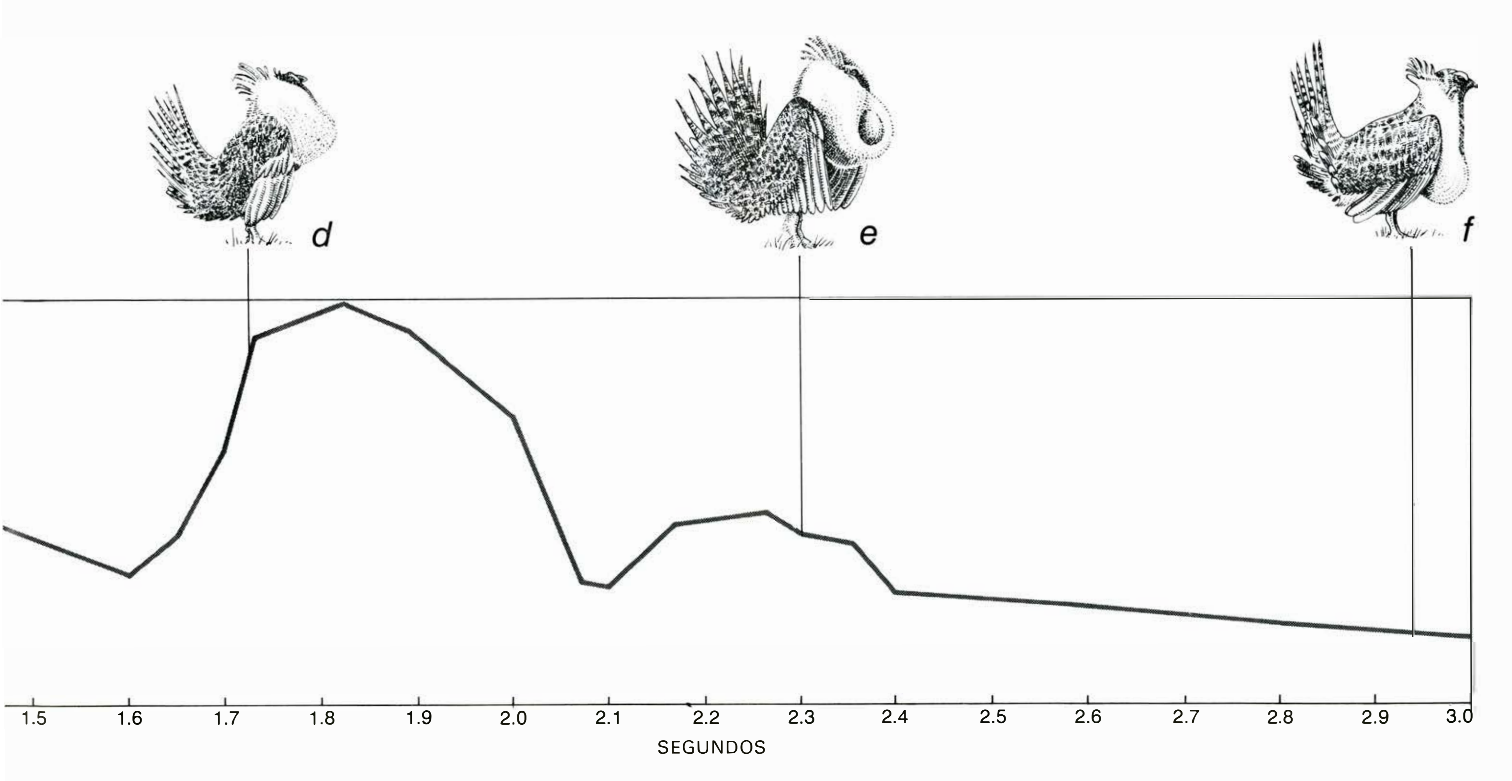
La hembra puede escoger un lugar de apareamiento de acuerdo con otras pautas que no sean la discriminación entre las diferencias de comportamiento de los machos. Por ejemplo, los territorios de los machos son menores cerca del centro de cópula que en la periferia del lek, de modo que la congregación de machos es más densa en la proximidad del centro. Las hembras podrían reconocer esta diferencia de densidad y elegir para aparearse aquellas zonas que tuvieran una mayor concentración de machos en exhibición. La información basada en la densidad, aunque no suficientemente específica para localizar de manera exacta el centro de cópulas identificaría su situación general.

Hay otra posible explicación que peca de mayor ambigüedad todavía. Por lo general, un centro de cópula suele ubicarse en un punto en el interior del lek en el que la vegetación de artemisa es dispersa. Suele haber varios de estos puntos comparativamente desnudos en el interior del lek, pero sólo uno corresponde a la localización de un centro de cópula.

Una señal muy específica de la localización de un centro de cópula es la presencia de una densa banda de hembras. Sin embargo, esta información sólo sería útil a aquellas hembras que llegaran al lek más entrada la mañana que el resto.

He podido observar que las hembras que llegan a un lek tienden a seguirse una a otra y a detenerse en grupo aunque se hallen todavía a cierta distancia del centro de apareamiento. Esta atención de las hembras al comportamiento mutuo permite, quizás, a las primerizas seguir a las más experimentadas hasta el centro de cópula, pero sólo en el caso de que las más experimentadas ya hubieran aprendido la situación del centro a principios de la estación de reproducción, o bien recordaran su situación de un año anterior. Una característica de las hembras de un año de edad, señalada por Robert Eng, de la Universidad estatal de Montana, y por Paul Dalke, de la Universidad de Idaho, y sus colegas, constituye un mecanismo por el que las hembras experimentadas podrían muy bien informar a las primerizas. Eng y Dalke observaron que las hembras primerizas llegan más tarde al lek y ponen sus huevos más avanzada la estación que las hembras del segundo año y más viejas. En otras palabras, las hembras primerizas disponen del asesoramiento de las experimentadas.

Todavía no se sabe exactamente qué combinación de señales guía a la hembra, a la hora de seleccionar un lugar para aparearse. Pero hay dos hechos (el que las hembras se sigan unas a otras en su camino hacia el lek y el que se dispongan en un grupo más denso en el centro de cópula) que sugieren que las interacciones entre las propias hembras, y no meramente la respuesta de éstas a las exhibiciones de los machos, controlan gran parte de su comportamiento reproductor, si no todo.



va la postura cerca del máximo. La sacudida hacia arriba que acompaña al hinchamiento del saco se advierte en *b* y *d*; la contracción, que produce un sonido similar a un chasquido, se advierte en la secuencia *e*, unos 2,3 segundos después de haber comenzado el gallo macho su exhibición.

La pregunta clave en este caso no es "¿cómo eligen las hembras un macho determinado para copular?", sino "¿cómo adquieren su territorio los machos en un centro de cópula?" La continuidad y persistencia de un centro de cópula en la misma situación dentro del lek, año tras año, no puede reflejar una preferencia duradera de la población de hembras por determinados machos. Esto resulta evidente al considerar que la mortalidad en la población masculina representa alrededor del 50 por ciento anual. Para reconocer la manera en que algunos machos alcanzan el centro de cópula debemos analizar la dinámica social que regula el acceso a los territorios.

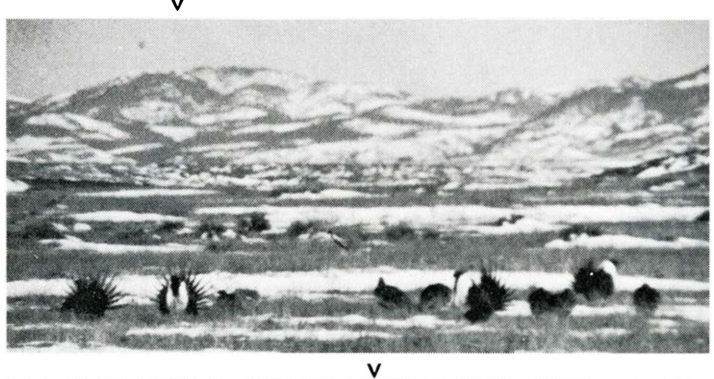
En los leks que observé, el proceso de acceso puede describirse como un llenado de vacíos que está orientado centrípetamente. Siempre que un gallo macho desaparecía durante la noche, su lugar era ocupado por una o más aves vecinas cuyo territorio se hallaba más alejado del centro de cópula. Nunca observé que un macho ocupara un puesto vacío que se hallara más lejos del centro que su propio territorio.

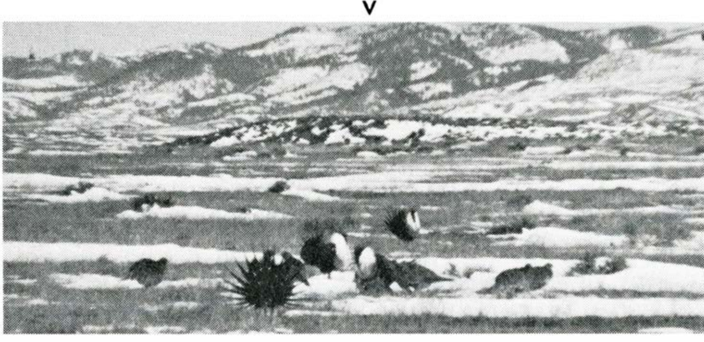
El proceso de ocupación de lugares vacíos era gradual. Frecuentemente, el primer día ninguno de los vecinos del macho desaparecido realizaba ninguna intrusión de importancia en el territorio abandonado. Sin embargo, al día siguiente o en los inmediatos, un vecino periférico extendía sus actividades de pavoneo cada vez más lejos, hasta que convertía el lugar vacío en su nuevo territorio. El movimiento centrípeto dejaba desde luego vacío el antiguo territorio del recién llegado, y ello iniciaba un desplazamiento centrípeto similar por parte del vecino periférico más cercano.

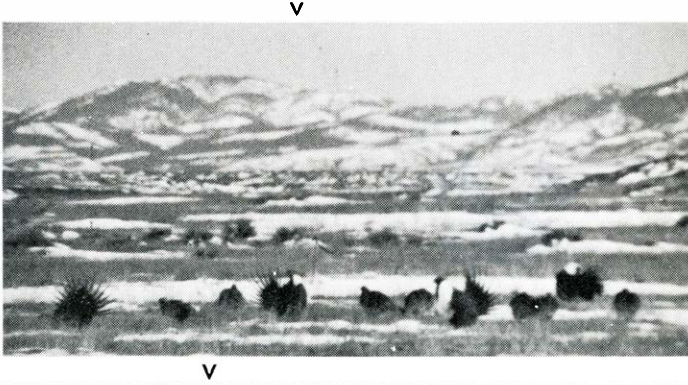
No siempre puede predecirse el modelo de una manera exacta, a pesar de que ningún macho se mueva nunca en dirección centrífuga ni salte sobre un territorio intermedio. Así, dos vecinos equidistantes del centro dividen a veces un lugar vacío más céntrico entre ambos, pero en otras ocasiones un ave reclama por entero el lugar vacío más central y la otra ignora el lugar vacío nuevamente creado y permanece en su viejo territorio. Dejando aparte estos detalles, en cualquier época de apareamiento los machos tienden a moverse gradualmente cada vez más cerca del centro de cópula, ocupando simplemente los lugares vacíos.

Varias razones hacen probable que el gallo macho continúe su avance hacia el centro de cópula, un año tras otro. En primer lugar, en muchas especies de tetraónidos que forman leks los machos mantienen un contacto, por lo menos superficial, con sus leks durante todo el año. En los días cálidos de otoño e invierno visitan brevemente sus leks por la mañana. No suelen mostrarse muy aparatosos en estas ocasiones fuera de temporada, pero las visitas pueden permitir que determinados machos mantengan contacto con sus vecinos de lek durante todo el año e incluso de un año para el otro. A este respecto, ciertos estudios de agregación han demostrado que, en cuanto el gallo macho ha establecido su territorio en un lek, raramente se desplaza hacia otro lek en épocas de reproducción posteriores.

Las observaciones de la actividad en el lek registradas a finales de invierno, justo antes de que los machos empiecen a congregarse con regularidad, revelan la falta de competencia vigorosa por las posiciones centrales. En una mañana determinada aparecen sólo unos cuantos machos de los que se hallan establecidos en el lek, y la actividad de las aves es bastante escasa. Parece razonable, pues, suponer que la ocupación de los lugares que se hallan vacíos a causa de la mortalidad invernal es un proceso gradual que comienza con el retorno de los machos sobrevivientes a sus posiciones aproximadas primitivas. Así ocurre efectivamente en el gallo de cola afilada *(Pedio-*









LA ACTIVIDAD QUE SE DESARROLLA en un centro de cópula durante un período de 25 minutos a mediados de abril, en el lek del Ford's Creek, en Montana, puede seguirse en los ocho fotogramas de una película de filmación diferida reproducidos aquí. El centro de cópula coincidía con el territorio de un gallo macho; la V sobre cada imagen sirve para localizar al macho. En el primer fotograma pueden verse unas siete hembras y cuatro machos; las hembras se hallan congregadas en el interior del centro de cópula *(centro)*, y el macho *A* *(V)* acaba de dirigirse hacia uno de los bordes de su territorio para enfrentarse a uno de los tres machos vecinos. El encuentro duró menos de cuatro segundos; en la imagen siguiente, el macho *A* ha vuelto al grupo que forman las hembras. En la tercera fotografía, unos cinco minutos más tarde, el macho *A* se encuentra en el lado opuesto de su territorio, donde una hembra solicita la cópula. En la cuarta imagen, menos de un minuto des-

ecetes phasianellus), como ha podido documentar con detalle Henry Kermott, de la Universidad de Minnesota. Este investigador encontró que el proceso de movimiento gradual hacia el centro de cópula empieza en la primera estación reproductora del macho y continúa de un año para otro.

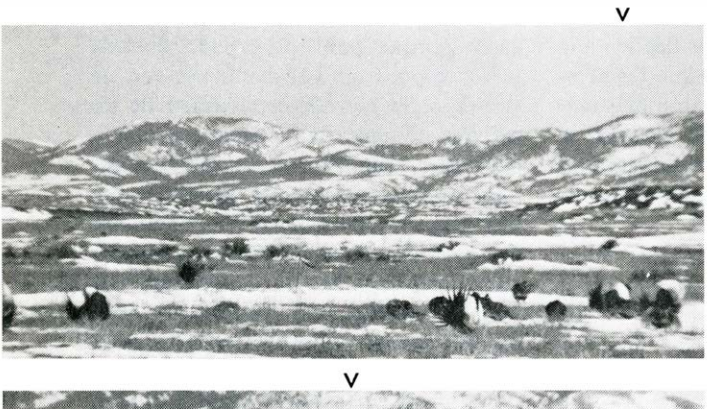
Los machos de gallo de las artemisas de un año, fácilmente identificables porque sus plumas caudales son menos aguzadas que las de los machos más viejos, llegan al lek mucho más avanzada la época de reproducción que éstos. Los individuos jóvenes visitan probablemente varios leks antes de instalarse en uno de ellos. La mayoría de machos de un año de edad no han establecido todavía sus propios territorios en el lek cuando hacen acto de presencia las hembras. A mediados de abril, que es la época de apareamiento, los machos jóvenes toman posiciones en la periferia del lek. A su debido tiempo cumplen el rito territorial normal, retornando regularmente cada mañana para recuperar su dominio elegido. Aquí la suerte del sorteo puede desempeñar una parte importante en el éxito futuro de los jóvenes machos. Los leks rara vez son exactamente simétricos; en consecuencia, algunos de los machos de un año encontrarán menos territorios ocupados entre ellos y el centro de cópula que otros. Asimismo, la ventaja de una posición inicial superior puede quedar anulada si la muerte aleatoria de las aves más viejas no deja vacantes los territorios intermedios.

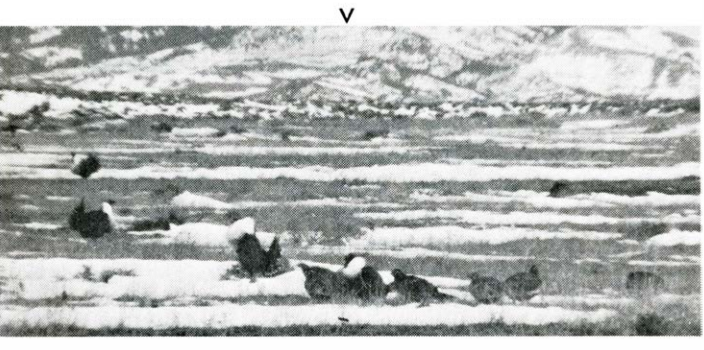
Las pruebas observacionales permiten configurar una hipótesis sobre el éxito sexual de los machos en los siguientes términos. En su primer año, el gallo de las artemisas establece un territorio en borde de un lek. En años sucesivos el ave tiende a volver al mismo lek y también a la misma situación general dentro del lek. La localización del ave cambia de forma centrípeta, tanto a lo largo de una estación como de un año para otro, y su camino al centro de cópula (y al éxito reproductor) es siempre función de la probabilidad. Aunque sin duda existe gran variedad en la edad de los machos que alcanzan el centro de cópula, los machos de un año de edad no suelen aparearse nunca, y los de dos años tienen menos probabilidades de conseguir copular que los más viejos. Existe, pues, una trayectoria ontogenética regular, una progresión en el desarrollo, mediante la cual un macho llega al éxito reproductor completo. En términos más sencillos, la clave del éxito no es la belleza sino la edad.

Lo que se ha dicho hasta aquí es aplicable a los machos. ¿Existe una trayectoria ontogenética de este tipo para las hembras de gallo de las artemisas? Evidentemente no. Casi todas las hembras se aparean cada estación; empiezan a reproducirse al primer año y ponen anualmente el mismo número aproximado de huevos mientras sobreviven. El joven gallo de las artemisas, por otra parte, no sólo se reproduce raramente, sino que además está menos desarrollado que los machos más viejos en cuanto a su morfología y a su fisiología. Su plumaje no está tan desarrollado como el de los machos mayores, e incluso en plena época de reproducción sus testículos son menores.

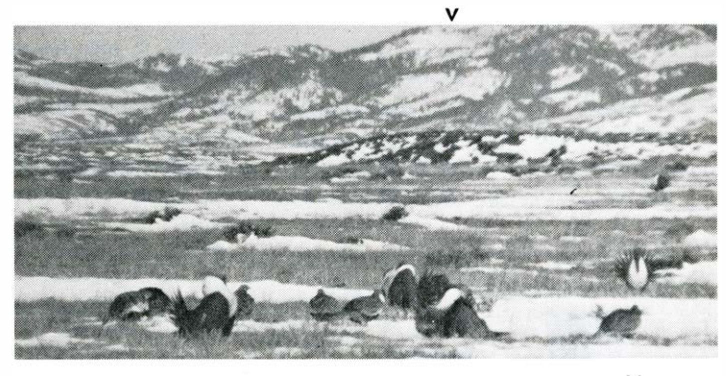
He llamado "bimaturismo" sexual a esta diferencia en los ciclos biológicos del macho y de la hembra del gallo de las artemisas, que representa un caso análogo al dimorfismo sexual. El bimaturismo no es sólo característico del gallo de las artemisas y de otras aves que se exhiben en leks, sino de la mayoría, por no decir todos, de los animales que presentan un sistema de apareamiento poligínico. Ni siquiera la sociedad humana se halla exenta: donde se practica el matrimonio poligínico la diferencia de edad entre el macho (más viejo) y la hembra (más joven) en el primer matrimonio es sustancial.

Hasta aquí he intentado analizar los





pués, el macho *A* copula por primera vez *(primer plano hecho con zoom).* Unos tres minutos más tarde *(quinta imagen)* el macho *A* se aparea con una segunda hembra; los machos vecinos continúan exhibiéndose. No han pasado dos minutos *(sexta imagen)* cuando los machos vecinos se han dispersado por las zonas más alejadas de sus territorios y sólo uno de los tres *(derecha, atrás)* aparece en la imagen. El macho *A* se pavonea solo entre el grupo de hembras.





Unos cuatro minutos después *(séptima imagen)* el macho *A* copula por tercera vez. El ave se apareó cinco veces más en los ocho minutos siguientes, fecundando así a todas las nueve hembras menos una, que se habían reunido en el centro de cópula por la mañana. Hacia el minuto nueve *(octavo fotograma)* el macho *A* se ha dirigido a la derecha para retar al macho que ocupa este territorio adyacente. El grupo de hembras que habían acudido al lek en esta época de apareamiento se está dispersando, como puede observarse.

mecanismos del comportamiento que regulan la estructura del sistema de lek del gallo de las artemisas. ¿Cómo pudo haber evolucionado un sistema similar? Una característica del sistema (la congregación de machos que se exhiben) pudo haber evolucionado, por lo menos entre los tetraónidos, de una forma directa. Agruparse en la época del apareamiento debe tener un valor adaptativo para los tetraónidos en los hábitat abiertos, porque protege a los machos alardeantes de los ataques por sorpresa de un depredador a la mortecina luz del amanecer. Los tetraónidos poligínicos que viven en bosques no forman leks: cada macho ocupa por separado un gran territorio y realiza sus exhibiciones en lugares muy distanciados. En terreno abierto, sin embargo, los machos de todos los tetraónidos poligínicos se exhiben en grupo. La presencia de tantos centinelas, por así llamarlos, ofrece a cada macho cierta protección frente a ciertos depredadores (águilas y coyotes). Yo mismo he sido testigo de varias escapadas por los pelos de águilas atacantes que parecían conocer la localización del lek. Los depredadores, volando a ras del suelo, utilizaban para ocultarse el resguardo de unas colinas bajas, y con frecuencia se acercaban sin ser observados hasta unos 100 metros del lek antes de su arremetida final.

Lo que resulta más difícil de explicar es la evolución de una estructura social que presenta una distribución tan notoriamente desigual de las cópulas entre los machos: los gallos de las artemisas practican la poliginia más extrema que se conoce en todas las aves. La mayoría de biólogos están hoy día de acuerdo en que la evolución del comportamiento social, como la evolución de cualquier otra característica, se explica mejor en términos de la optimación de la "eficacia darwiniana" del individuo, es decir, la optimación de la tasa a la que los genes del individuo son transferidos a sus descendientes. Una explicación de la evolución de un sistema de apareamiento debe dar cuenta de las ventajas del mismo con respecto a la eficacia darwiniana de los machos y de las hembras.

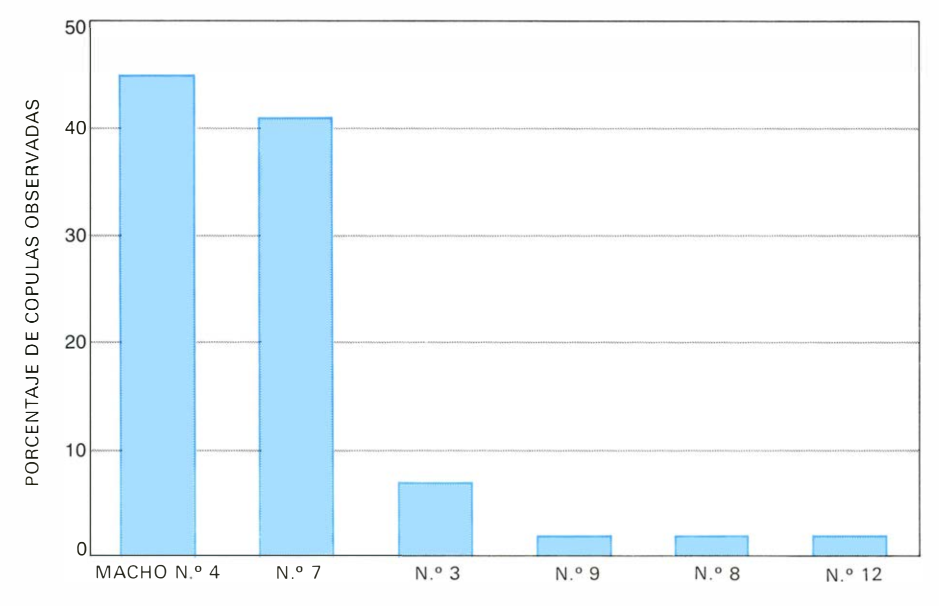
En lo que respecta a la eficacia de los machos en las sociedades poligínicas, el retraso que se produce hasta que éstos consiguen una reproducción favorable complica las cosas. La tasa a la que los genes de un individuo se propagan a sus descendientes no depende sólo de la fecundidad del individuo, sino también de la edad a la que éste y sus descendientes empiezan a reproducirse. En términos más sencillos: un retardo en la reproducción aumenta el tiempo de generación y reduce así la tasa a la que se multiplican los descendientes. Un macho que retarda la reproducción tiende de hecho a sacrificar las ventajas de una fecundidad elevada una vez ha empezado la reproducción.

Para comprender la manera en que esto ocurre, considérese una sociedad aviar imaginaria en la que machos y hembras tengan idénticas tasas de supervivencia, si bien los machos lleguen a la madurez sexual más tarde que las hembras. En cada cohorte morirán más machos que hembras antes de alcanzar la edad reproductora; de ello resulta que en cada estación se reproducirán menos machos que hembras. Una sociedad de este tipo está abocada necesariamente a la poliginia.

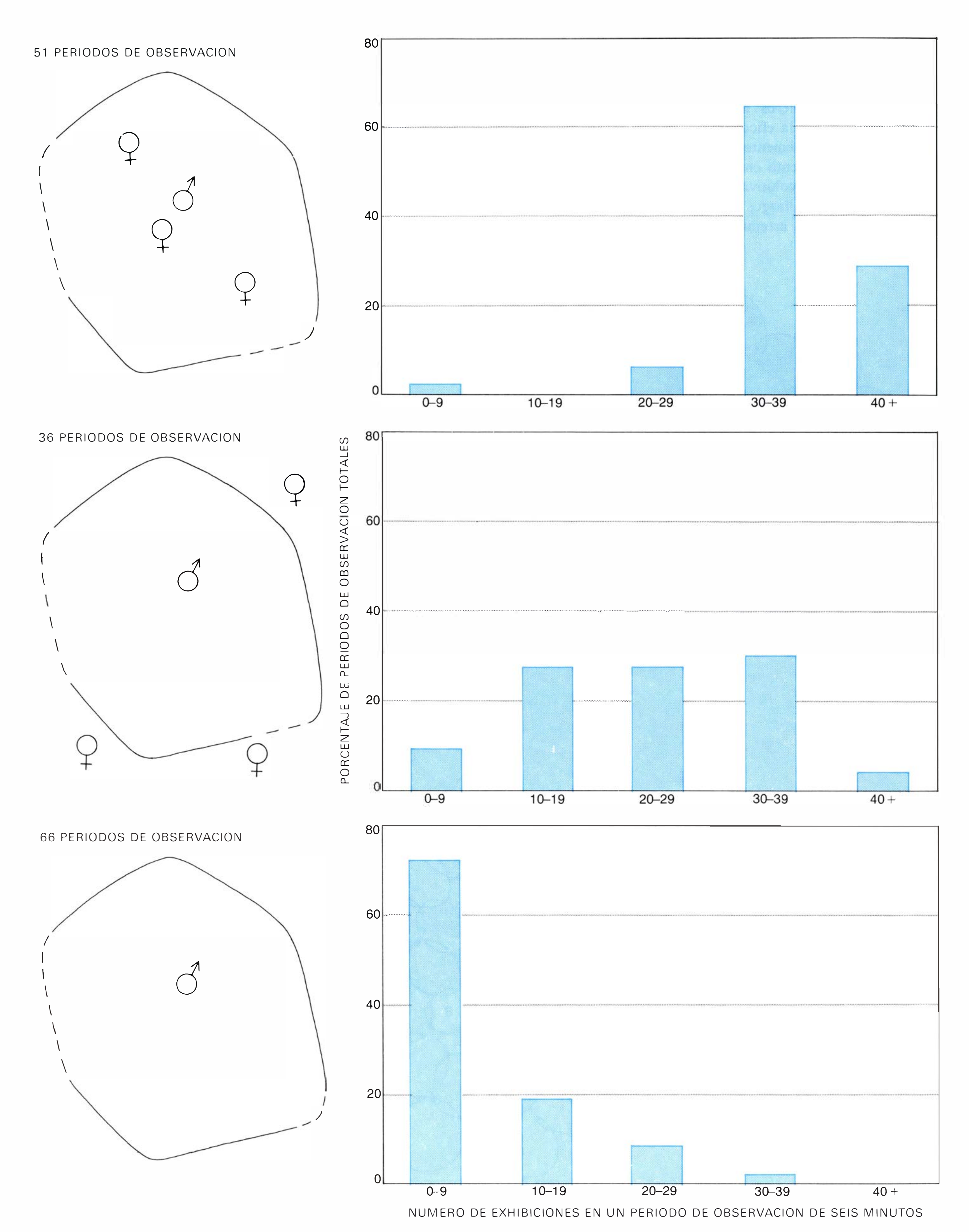
Imagínese además que el tamaño de la población es constante, de modo que a lo largo de su vida cada ave deje por término medio un descendiente de su propio sexo, y también que una vez los machos empiecen a reproducirse compartan todos ellos por igual a las hembras disponibles. En una sociedad tan simple la poliginia no ofrecerá ventajas ni desventajas a los machos. Cuanto más tiempo difieran éstos la reproducción, tanto más fecundos serán cuando alcancen la madurez sexual. Los dos efectos se compensan entre sí de manera perfecta. Mas para que el carácter de reproducción diferida pueda evolucionar, el macho de tetraónido debe obtener alguna ventaja en la propagación de sus descendientes.

Durante mucho tiempo han gozado de crédito dos teorías inadecuadas y muy simplistas sobre las ventajas de los machos en los sistemas de apareamiento de lek. Puede llamárselas, respectivamente, la teoría de la "atracción sexual" y la teoría del "juicio de Dios". La primera teoría propugna que los machos que tienen éxito en las sociedades poligínicas poseen una eficacia evolutiva superior a la de los machos que no se aparean, y ello debido a que tienen una fecundidad superior. La segunda teoría sostiene que los machos con éxito tienen mayor eficacia porque han sobrevivido durante más tiempo. Como acabamos de ver, sin embargo, la elevada fecundidad de los machos de gallo de las artemisas que se aparean no significa necesariamente que tales machos posean una eficacia evolutiva superior, y tampoco puede aducirse que esta última sea producto de una mayor duración de la vida. La eficacia consiste, en cambio, en la distribución óptima del tiempo y de la energía de cada uno de los gallos machos para la consecución de su supervivencia y para el ejercicio de su fecundidad a lo largo de toda su vida.

Parece plausible, cuando menos, una ventaja hipotética para los machos en las sociedades poligínicas. Si los machos que se reproducen menos tienen más probabilidades de sobrevivir, un tal aumento en la supervivencia, si es suficiente, podría aumentar la eficacia evolutiva de los machos que retrasasen la reproducción a una edad más avanzada. Pero no está claro todavía si una hipótesis basada en estas premisas puede ayudar a explicar la poliginia del gallo de las artemisas. La verificación en la naturaleza no deberá ser fácil: requeriría compa-



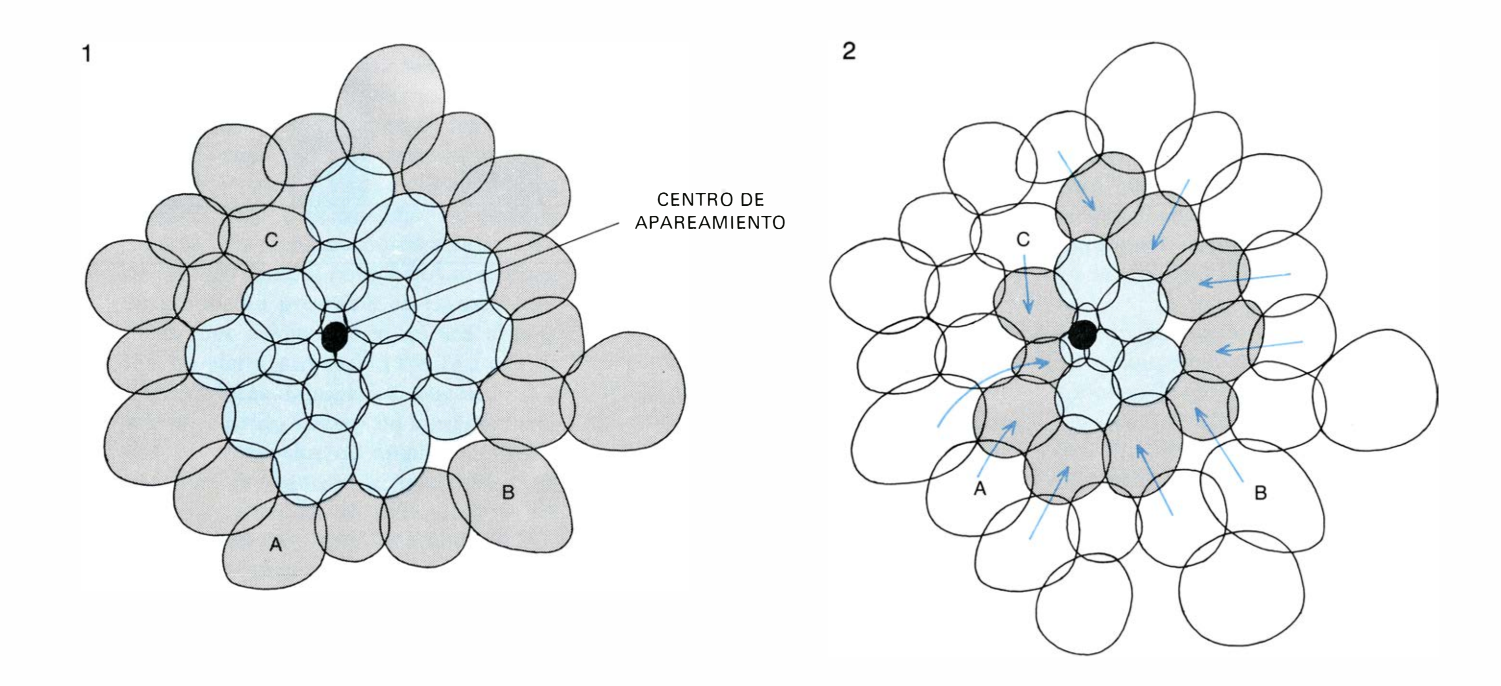
REGISTRO a lo largo de una estación de reproducción de los seis machos que se aparearon en el lek Dry Sandy; las dos aves más activas realizaron el 86 por ciento de las 42 cópulas observadas.

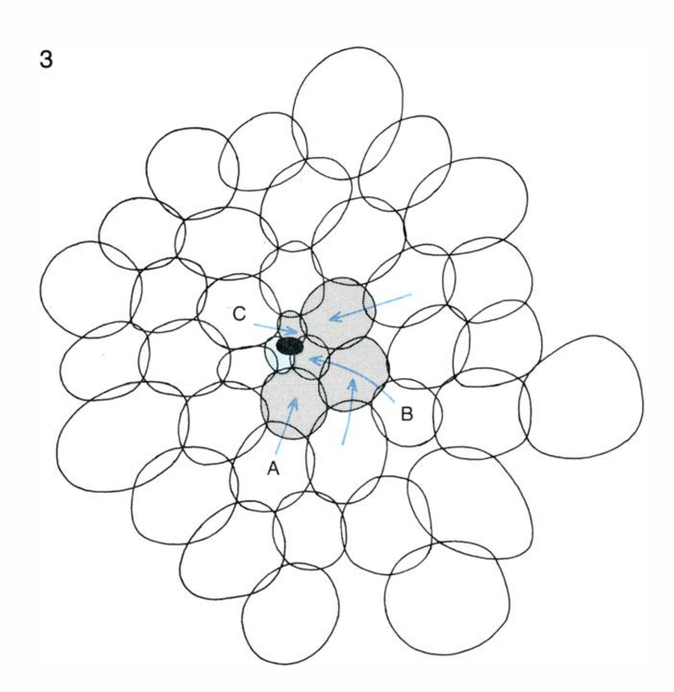


FRECUENCIA DE LAS EXHIBICIONES de los machos de gallo de las artemisas; está directamente relacionada con la proximidad de las hembras. Un total de 51 observaciones demostró que los machos se exhibían por lo general más de 30 veces durante un intervalo de 6 minutos, si las hembras se encontraban en el interior del territorio del macho *(arriba).* **Cuando las hembras se encontraban fuera del territorio del macho, aunque en el interior de un territorio vecino** *(centro),* **la frecuencia fue menor, si bien los machos todavía se pavoneaban de 10 a 39 veces cada 6 minutos. En caso de que no hubiera hembras en las cercanías** *(abajo)* **las exhibiciones bajaron a menos de 10 cada 6 minutos.**

raciones de la longevidad y de la fecundidad de los machos que empiecen a reproducirse a edades distintas.

Cuando los primeros naturalistas se interrogaron sobre la eficacia evolutiva, asimilaban frecuentemente la competencia de comportamiento entre individuos a la competencia evolutiva. De acuerdo con mis propios hallazgos en el estudio de los gallos de las artemisas, no puede aceptarse en ningún caso una tal identidad. La competencia del primer tipo tiene que ver con el individuo que gana los combates,. mientras que la competencia del segundo tipo se refiere a los genes que cada individuo transmite a sus descendientes. Después de observar un lek sería tentador sacar la conclusión de que los machos con éxito son simplemente los ganadores en la competencia con los machos que no logran aparearse. En un sentido de comportamiento así es: los machos del centro de cópula expulsan a los intrusos de su territorio con amenazas y agresiones declaradas. Esto, sin embargo, no es necesariamente competencia evolutiva. Si los machos aumentan de hecho su eficacia evolutiva retrasando la actividad reproductora total, entonces el hecho de que un macho





AVANCE HACIA EL CENTRO DE COPULA. En un principio *(1)*, 20 jóvenes machos han establecido territorios *(gris)* en la periferia de un lek durante su primera estación de apareamiento; las posiciones más centrales están ocupadas por machos mayores. Es pura casualidad que el territorio del macho *C* esté más cerca del centro de cópula *(negro)* que los territorios de los machos *A* y *B*. En la primavera siguiente *(2)* murieron cerca de la mitad de los machos de un año de edad del año anterior, y también la mitad más viejos. Los machos *A*, *B* y *C* y otros de su cohorte *(gris)* se hallan ahora más cerca del centro de cópula; la mejor posición de partida del macho *C* lo ha llevado más cerca que ninguno. Al año siguiente *(3)* sólo sobreviven cinco machos de la cohorte. Los territorios de los machos *B* y *C* se solapan ahora con el centro de cópula y el territorio del macho *A* se encuentra adyacente al mismo; las trayectorias ontogenéticas de los machos *B* y *C* significan que engendrarán muchas de las crías producidas durante la estación. El macho *A*, si sobrevive, alcanzará probablemente el centro de cópula al año siguiente.

más viejo excluya con éxito a otro más joven de la oportunidad de copular puede que no tenga nada que ver con diferencias en la eficacia darwiniana. Es muy posible que ambos gallos actúen siguiendo una estrategia evolutiva que aumenta la eficacia de ambos.

¿Ofrece alguna ventaja el sistema de apareamiento poligínico a la hembra de gallo de las artemisas? Una consecuencia del apareamiento de este tipo es que el cuidado paterno por parte de los machos se reduce, o no existe en absoluto. De las 16 especies de tetraónidos, 12 son poligínicas, y sólo una de las cuatro especies monógamas practica el cuidado paterno dual. Los cuidados paternos desempeñados por un solo progenitor serían ciertamente desventajosos, tanto para las especies monógamas como para las poligínicas, si los pollos de los tetraónidos necesitaran, como ocurre en muchas aves recién salidas del huevo, un período prolongado de alimentación por parte de los padres hasta que alcanzaran un estadio de independencia. Sin embargo, parecidamente a lo que sucede con las gallinas domésticas, las 16 especies de tetrónidos dan origen a crías precoces que son capaces de seguir a su madre y de alimentarse por sí mismas poco después de salir del huevo; de ahí que la desventaja potencial del apareamiento poligínico se ha reducido muchísimo. No obstante, tres especies de tetraónidos prestan a sus crías cuidados unipaternales, pero practican la monogamia. Evidentemente, la evolución de una pauta social de cuidados unipaternales, y por ende la evolución de crías precoces que hacen posible el cuidado unipaternal, aunque prerequisitos necesarios para la evolución de la poliginia, no constituyen factores causales suficientes.

En conjunto, la ventaja adaptativa que gana la hembra de cualquier especie de tetraónido depende de las ventajas de la estrategia del macho. Los genes de cada hembra son transmitidos por igual a sus descendientes masculinos y femeninos; en consecuencia, las mejores estrategias para cada sexo no se hallan en conflicto. Al aparearse con machos que practican la estrategia masculina óptima, las hembras aumentan su propia eficacia evolutiva: la tasa a la que sus genes se propagan a sus descendientes. Por tanto, y de generación en generación, tanto el macho como la hembra de gallo de las artemisas han desarrollado una relación de apareamiento en concordancia con un delicado equilibrio entre las desventajas de los cuidados unipaternales y las ventajas del bimaturismo sexual.

El principio de adyacencia en la percepción visual

El sistema visual integra la información exterior procedente de diversas fuentes. El principio de adyacencia describe cómo contribuyen las señales relativas a que se complete esa integración

Walter C. Gogel

En sus *Principios de psicología de la Gestalt",* se preguntaba Kurt Koffka "¿por qué, pues, parecen las cosas lo que parecen?" Y argüía que, seguramente, no "porque son lo que son". Si las cosas parecieran lo que parecen sólo por lo que son, seguirían pareciendo las mismas mientras siguieran siéndolo. Pero el hecho es que no siguen pareciendo las mismas: parecen diferentes según sea su contexto, y, más en especial, según se las vea aisladas o junto a otros objetos.

Hay una curiosa prueba de hasta qué punto influye el contexto en la percepción del movimiento: Si se mueve repetidamente, a derecha e izquierda, un único foco de luz en un campo obscuro, se percibe como haciendo precisamente esto, moverse horizontalmente. Pero si a la vez se mueve cerca otro foco hacia arriba y hacia abajo, los movimientos percibidos de los dos focos, que se mueven el uno respecto al otro en ángulo recto, pasan en la percepción a diferir marcadamente de sus reales movimientos horizontal y vertical *(véase A en la ilustración de la página 68).* Como lo ha expresado Gunnar Johansson, de la Universidad de Uppsala, los movimientos percibidos comprenden vectores de movimiento relativo (los focos parecen acercarse y alejarse uno de otro a lo largo de una senda diagonal) y vectores de movimiento común (puede parecer también que se mueven en grupo a lo largo de otra diagonal que corte en ángulo recto a la primera).

Este claro cambio en la percepción de un objeto al introducir otro objeto (u otros varios) en el campo de visión ha podido comprobarse no sólo en lo que respecta al movimiento, sino también en lo referente a características percibidas tales como el color, el tamaño, la forma, la distancia y la orientación. Los factores que determinan las características percibidas de un objeto independientemente de otros objetos son señales absolutas; los factores que hacen que cambie la percepción cuando entran en juego otros objetos son señales relativas. El objeto cuya característica percibida se está midiendo es el objeto-prueba *(test object);* un objeto adicional cuya presencia modifica la percepción es el objeto-inducción *(induction object).* Las investigaciones realizadas en mi laboratorio, en la Universidad de California en Santa Bárbara, han sugerido una regla general que ayuda a predecir el poder de una señal relativa para modificar una percepción. Esta regla, que nosotros llamamos el principio de adyacencia, establece que la importancia que el sistema visual otorga a una señal relativa está en proporción inversa a la separación aparente entre el objeto-prueba y el objeto-inducción en el espacio tridimensional. Aquí consideraré la prueba más vigorosa del principio de adyacencia en la percepción del tamaño, la orientación, la distancia y el movimiento, y a continuación trataré brevemente de lo que podrían ser algunas de las consecuencias de tal principio.

Un supuesto básico, que se sobreentiende al medir efectos de adyacencia y está bien corroborado por los datos experimentales, es el de que el sistema visual es capaz de combinar información cuyas fuentes no concuerdan. Sean, por ejemplo, dos fuentes, una que de suyo llevaría a una percepción, y otra que de suyo llevaría a otra percepción diferente de aquélla. El supuesto dice que cuando estas dos fuentes se presenten juntas originarán una tercera percepción a la que ambas contribuyen. Esta síntesis de informaciones diversas queda demostrada por el aparente movimiento que se induce en un punto mediante la oscilación de un marco *(véanse B y C en la ilustración de la página 68).* Si un punto físicamente estacionario es rodeado de un marco que se mueve reiteradamente de izquierda a derecha en un campo por lo demás obscuro, el punto parecerá moverse en dirección opuesta a la del movimiento físico del marco. En un caso más complicado, el punto "se mueve" verticalmente: hacia arriba cuando el marco va hacia la derecha, y hacia abajo cuando el marco va hacia la izquierda. Aquí el punto tendrá dos factores componentes de su aparente movimiento, siendo uno de esos factores una moción horizontal inducida y resultando el otro de las señales asociadas al movimiento vertical físico. Ambos factores se suman vectorialmente, y, en consecuencia, el punto parece moverse en diagonal más bien que horizontal o verticalmente. Si en vez del marco entero se mueven sólo dos puntos en sentido horizontal *(D en la ilustración),* percíbese un efecto semejante de inducción horizontal.

No está claro si el movimiento vertical aparente del punto dentro del marco es atribuible a una señal cinética absoluta (el movimiento vertical físico del punto) o a la señal de movimiento relativo entre el punto y los elementos horizontales del marco. El movimiento diagonal percibido podría ser la resolución de un conflicto entre señales absolutas (verticales) y relativas (horizontales), o bien la resolución de un conflicto entre diferentes señales relativas, una de las cuales implicara las partes horizontales del marco y la otra las partes verticales. Hay pruebas de que ambos tipos de conflicto entre señales son resueltos por el sistema visual, que puede combinar informaciones discrepantes, que provengan de señales

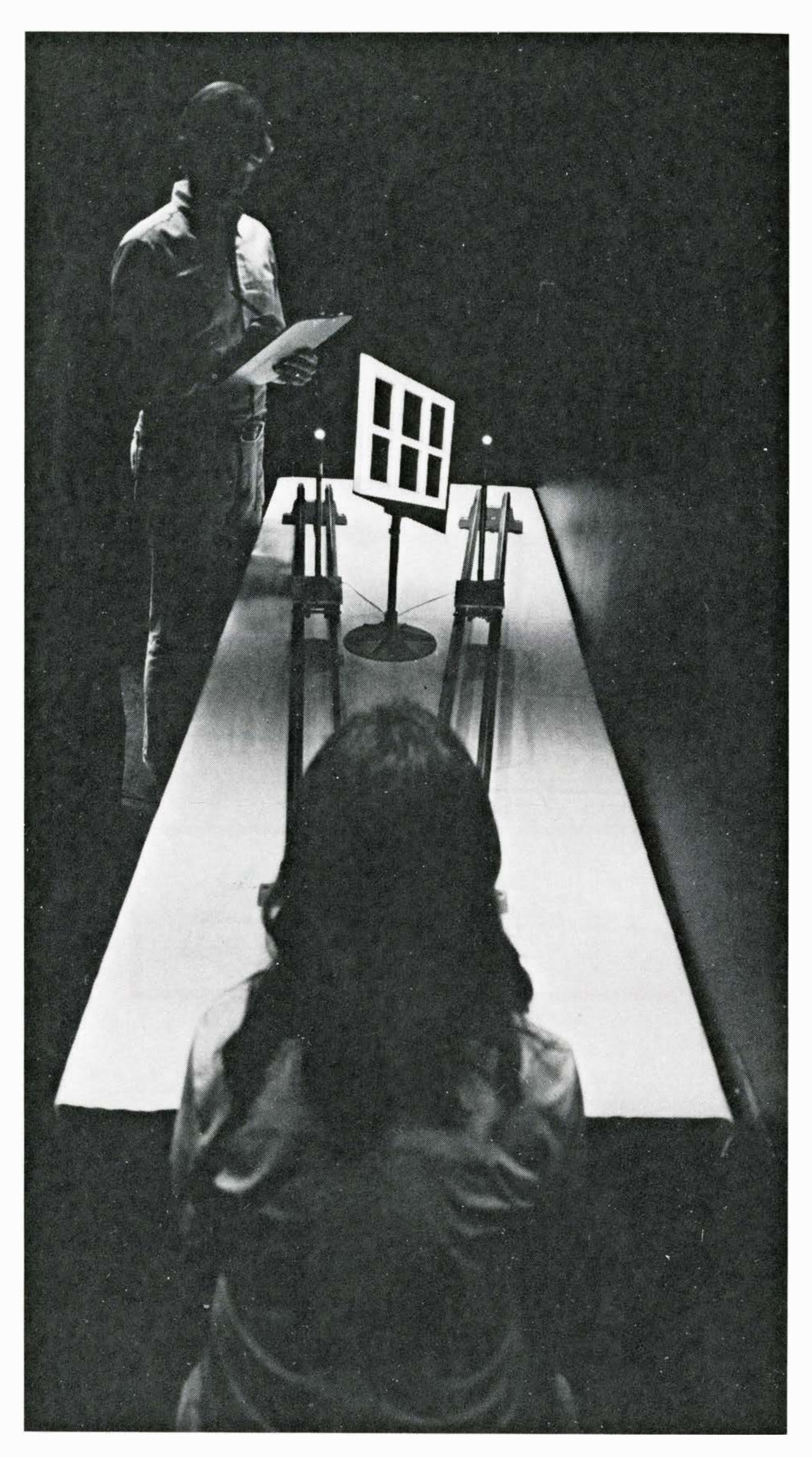
relativas o de señales absolutas o de unas y otras a la vez, transformándolas en una variedad de características "percibidas".

Un modo de medir efectos de adyacencia es variar la distancia que separa un objeto-prueba de un único objeto-inducción, en un plano frontal paralelo (es decir, paralelo al que contiene los ojos del observador, o perpendicular a la línea de visión) o en profundidad (a lo largo de la línea de visión). Según el principio de adyacencia, a cada aumento de la separación deberá disminuir la magnitud de la inducción. Otro modo consiste en cambiar la posición aparente del objeto-prueba respecto a dos objetos-inducción separados uno de otro en un plano frontal paralelo o en profundidad y que producen efectos opuestos. Según el principio de adyacencia, el efecto de inducción deberá cambiar de dirección a medida que el objeto-prueba se desplace desde la proximidad de un objeto-inducción hasta acercarse al otro.

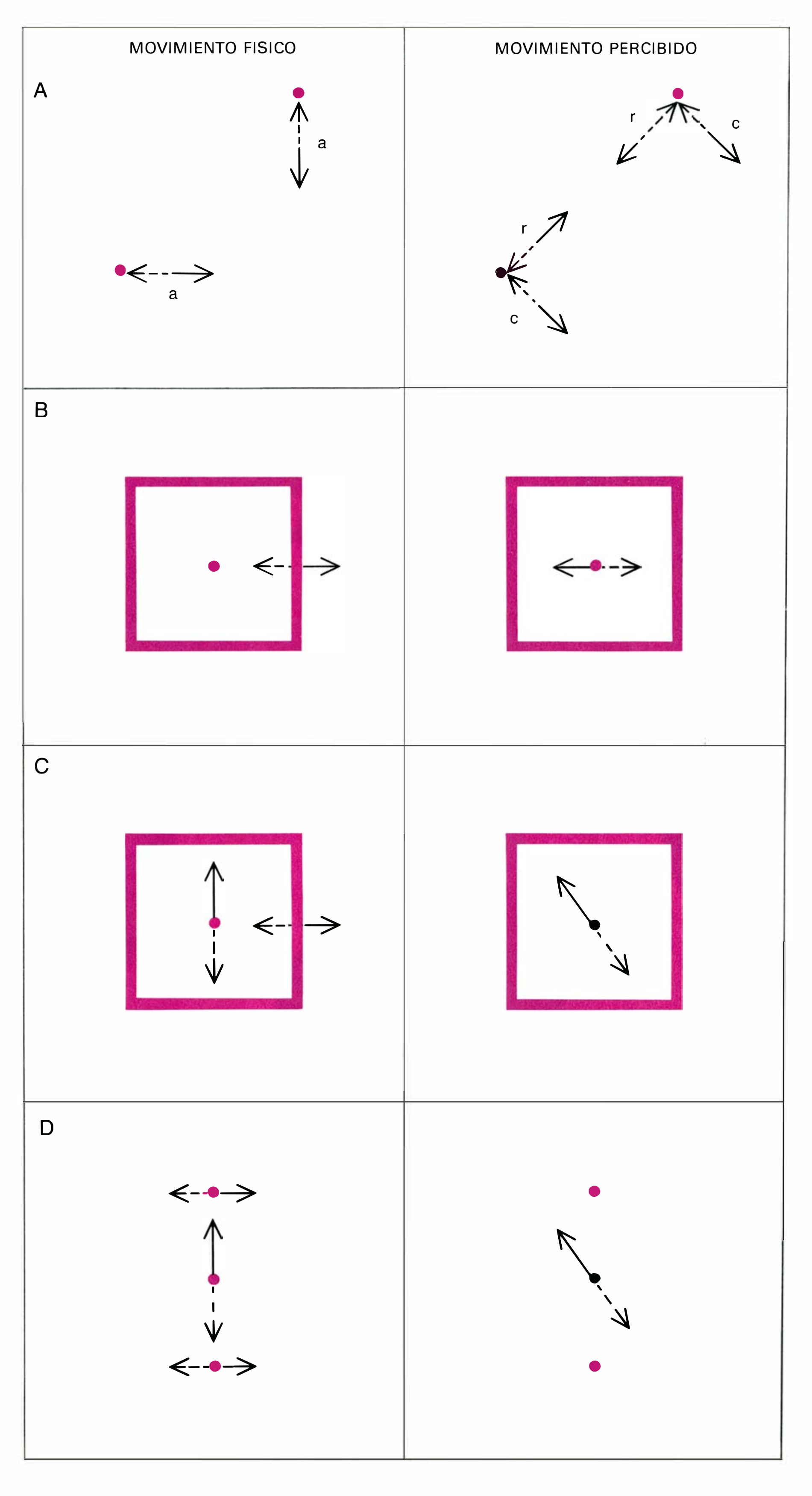
Empecemos por revisar nuestra prueba describiendo un experimento de este segundo tipo. A un observador se le presentan unos puntos de luz que se mueven en un campo obscuro: un objeto-prueba oscilando horizontalmente y dos objetos-inducción oscilando verticalmente *(véase la ilustración de la página 69)*.

La fase de los movimientos es tal que la moción aparente inducida en el objeto-prueba por un objeto-inducción está en ángulo recto con la moción aparente inducida por el otro objeto-inducción. En distintos momentos de la experimentación, el objeto-prueba se aparta progresivamente de uno de los objetos-inducción y se acerca al otro. Los observadores veían los puntos de luz con visión binocular, y se les pedía que fueran indicando, con ayuda de un puntero, el camino que, a su parecer, iba siguiendo en sus oscilaciones el objeto-prueba.

Los resultados corroboraron fuertemente el principio de adyacencia. El objeto-prueba *(2 en la ilustración de la página 69)* parecía moverse entre más arriba a la derecha y más abajo a la izquierda cuando estaba cerca del objeto-inducción *3 (a)*, aproximadamente en horizontal cuando estaba más o menos a la misma distancia de *1* y *3 (b)* y entre más abajo a la derecha y más arriba a la izquierda cuando estaba cerca del objeto-inducción *1 (c)*; es decir, que era más afectado por el objeto-inducción al que se hallaba más próximo, y era casi igualmente afectado por cada objeto-inducción cuando se hallaba equidistante de ellos. La magnitud promedio de la desviación del movimiento del punto 2



LOS EFECTOS DE ADYACENCIA ESPACIAL quedan comprobados al mirar con visión binocular una ventana trapezoidal, hecha con porciones ensombrecidas y enmascaradas de un papel electroluminiscente, y dos discos luminosos, todo ello en un campo visual obscuro. Un extremo de la ventana *(aquí el derecho)* está más cerca del observador que el otro extremo, pero perceptivamente está más lejos. El error de percepción se transmite a los discos. Estos se hallan físicamente equidistantes del observador, pero perceptivamente el disco derecho está más alejado que el izquierdo. Si se desplazan los dos discos a la vez, a lo largo de los carriles, parecen moverse a distintas velocidades mientras están cerca de la ventana. La influencia de la ventana en la posición (y en el movimiento) percibidos de los discos, disminuye a medida que éstos se van alejando de la ventana. (El sesgo hacia abajo con que aparece aquí la ventana se debe al ángulo de enfoque.) La aplicación del principio de adyacencia a estos fenómenos se ilustra en el diagrama de la página 72.



EL MOVIMIENTO APARENTE se induce en un objeto por otro objeto próximo. En el primer ejemplo *(A)* oscilan dos puntos, uno vertical y otro horizontalmente. Si se los presentara por separado, sólo uno cada vez, se percibirían moviéndose precisamente así. Cuando se presentan los dos a la vez, se perciben sus movimientos relativos en diagonal *(r)* y a menudo, se perciben sus movimientos comunes *(c)*, pero no sus movimientos absolutos *(a)*. Un marco oscilante *(B)* induce un movimiento horizontal aparente de fase opuesta en un punto físicamente estacionario situado dentro del marco. Si el punto se mueve físicamente arriba y abajo *(C)*, el marco produce un vector horizontal aparente, que, combinándose con el vector vertical, da por resultado un movimiento aparente del punto en sentido diagonal. Se induce una percepción similar cuando actúan, asimismo, dos puntos *(D)*.

con respecto a la horizontal fue de 17 grados en un caso extremo *(a)* y de 25 grados en el otro *(c)*, lo cual es considerablemente menos que los 45 grados que serían de esperar si sólo hubiese producido efecto la señal relativa originada por el objeto-inducción próximo. Está claro que, además del efecto dominante producido por el objeto-inducción próximo, hubo también algún efecto del objeto-inducción más distante o del movimiento absoluto del objeto-prueba, o de ambos.

Adaptamos este experimento con puntos no fijos a la situación en que sólo se presentaba un objeto-inducción, variando las distancias desde el objeto-prueba, presentando el punto *2* o con el *1* o con el *3*, en vez de hacerlo con ambos a la vez. De acuerdo con el principio de adyacencia, la magnitud del efecto de inducción tendió a disminuir a medida que el punto-prueba se separaba más y más del único punto-inducción. Aunque los resultados no fueron tan concluyentes como con los dos puntos de inducción, quedó probado que las señales absolutas producían algún efecto aun cuando el punto-prueba y el punto-inducción estuvieran adyacentes entre sí. En resumidas cuentas, los resultados de este experimento abonan claramente la conclusión de que la eficacia de las señales relativas disminuye conforme aumenta la distancia entre el objeto-prueba y el objeto-inducción; pero los resultados sugieren también que el objeto-prueba es influido un poco por las señales relativas procedentes del objeto-inducción más distante y por las absolutas procedentes del mismo objeto-prueba.

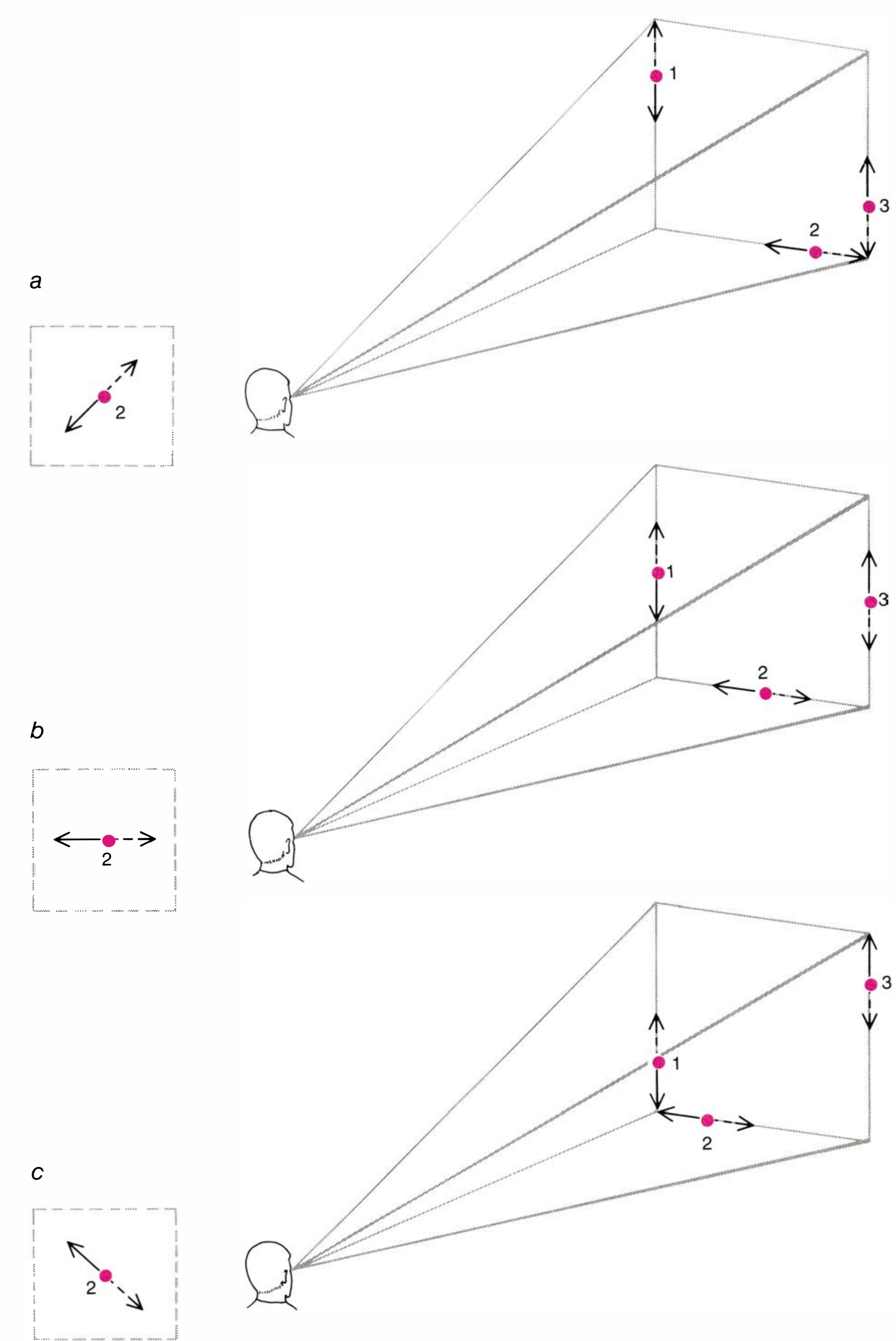
En otra serie de experimentos estudiamos el efecto que produce la separación frontal paralela sobre una característica diferente: la localización percibida en la distancia. La señal de distancia, objeto de nuestra investigación, era la proporcionada por el tamaño relativo que tienen en la retina del ojo objetos similares que un observador tomaría probablemente por idénticos de tamaño en la realidad física. Si dos de tales objetos cobran en la retina diferentes tamaños, el objeto retinianamente más pequeño es percibido como hallándose más alejado; la igualdad del tamaño retiniano de objetos similares es aceptada como información de que los objetos están a la misma distancia. Tres naipes y un cuadrado gris fueron presentados en forma de transparencias fotográficas retroiluminadas, todos ellos a la misma distancia, en una habitación obscura, y el sujeto las veía monocularmente. Un naipe era mayor y los otros dos, corrientes.

Si no hubiera estado presente el cuadrado gris, tapando parte del naipe ma-

yor y tapada su misma parte superior por uno de los naipes normales, las distancias relativas aparentes entre estos dos naipes (los objetos-inducción) habrían sido determinadas por las diferencias entre sus tamaños retinianos, y el de tamaño normal habría parecido estar más alejado que el grande. Pero el cuadrado gris proporciona una "señal interpositiva", que hace que el naipe normal esté delante de –y así más cerca que– el naipe de tamaño mayor. El objeto-prueba era el otro naipe de tamaño normal puesto aparte y más abajo, y fue presentado más cerca del naipe grande en un experimento *(a)* e igualmente separado del naipe grande y del de tamaño normal en el otro experimento *(b)*.

En ambos experimentos hay señales de tamaño relativas entre el naipe-prueba y cada uno de los naipes-inducción. La señal relativa entre el naipe-prueba y el naipe grande tiende a hacer que el naipe-prueba parezca estar detrás del naipe grande (y, por tanto, a causa de la señal interpositiva, muy por detrás del naipe de tamaño normal que aparece en la parte superior). La señal relativa entre el naipe-prueba y el naipe de más arriba tiende a hacer que el naipe-prueba parezca hallarse a la misma distancia que el naipe de más arriba (y, por tanto, a causa de la señal interpositiva, no detrás del naipe grande sino delante de él). O sea, que la presencia de la señal interpositiva indica que las dos señales de tamaño relativas que afectan al naipe-prueba están en conflicto.

El principio de adyacencia predice que, cuando el naipe-prueba está cerca del naipe-inducción más grande *(a)*, la señal de tamaño relativa entre el naipe-prueba y el naipe grande deberá hacer que el naipe-prueba parezca estar bastante detrás del naipe de tamaño normal que se encuentra más arriba. Y esto es lo que les pareció a nuestros sujetos de experimentación. Con el naipe-prueba igualmente apartado de ambos naipes-inducción *(b)*, el principio de adyacencia predice que será menor el efecto del naipe-inducción grande sobre el naipe prueba; y así sucedió: se dijo que el naipe-prueba parecía estar aproximadamente a la mitad de distancia, en profundidad, entre los dos naipes-inducción. Los resultados sugieren que las señales de tamaño relativas, puestas entre los naipes-inducción juntos o separados, contribuían a la aparente distanciación del naipe-prueba. Sin embargo, según el principio de adyacencia, la contribución era mayor para el naipe-inducción que estaba más cerca del naipe-prueba en el plano frontal paralelo.



EFECTO DE ADYACENCIA sobre señales de movimiento relativo. Se mide poniendo un objeto-prueba *(2)* a distintas distancias de objetos inductores opuestos *(1 y 3)*. Unos puntos de luz oscilan recorriendo tramos definidos por la longitud, dirección y fase *(línea continua o interrumpida)* de flechas, a la derecha; el punto *2* oscila cerca del punto *3 (a)*, a medio camino entre los puntos de inducción *(b)* o cerca del punto *1 (c)*. Las flechas de la izquierda indican cuál sería en cada caso la dirección del movimiento del punto *2* si el punto inductor más próximo determinase por completo la percepción del observador. De acuerdo con el principio de adyacencia, los resultados experimentales indican que, aunque probablemente los dos puntos inductores contribuyen al movimiento aparente del punto *2*, el punto más próximo ejerce mayor influencia. (En este experimento y en el que sigue sólo el objeto-prueba y los objetos-inducción eran visibles al observador.)

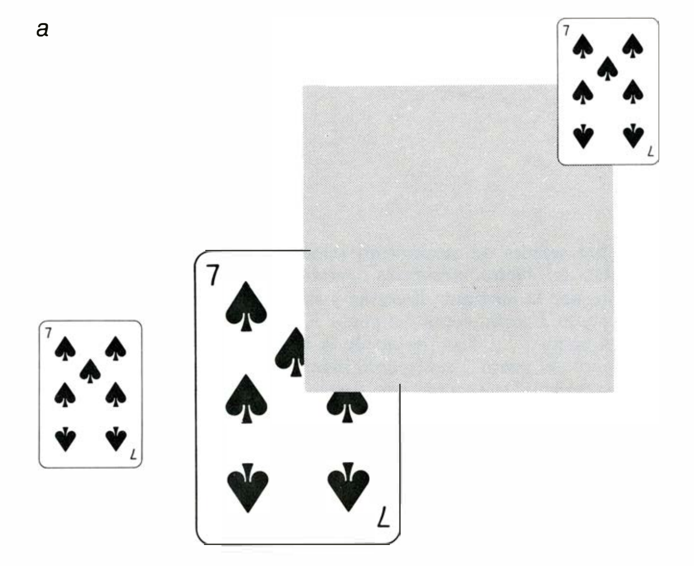
Dado que la separación en un plano frontal paralelo y la separación en la retina son en esencia proporcionales, pueden interpretarse los efectos de la adyacencia frontal paralela en términos de adyacencia en la retina. Que los efectos de adyacencia no vienen determinados por la separación retiniana lo prueban estudios que han demostrado que la eficacia de las señales relativas queda menguada por la separación, en profundidad, del objeto-prueba y los objetos-inducción. Por ejemplo, en algunos de nuestros experimentos, se presentan binocularmente dos objetos inductores calculados para que produzcan efectos opuestos sobre un objeto-prueba a diferentes distancias del observador; el objeto-

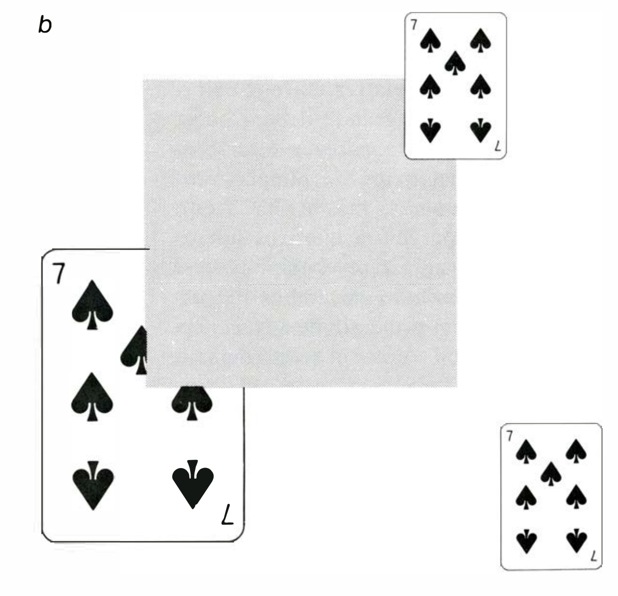
prueba puede presentarse a la distancia o bien del objeto-inducción más alejado, o bien del objeto-inducción más próximo, o en medio de los dos. En cada caso, se procura que el objeto-inducción más apartado sea de un tamaño lo suficientemente mayor que el más cercano como para que ambos cobren igual proporción en la retina del observador. Por lo que atañe, pues, a la estimulación retiniana, los efectos de estos dos iguales, aunque opuestos, objetos-inducción sobre el objeto-prueba deberían contrarrestarse y anularse. Mas, por otro lado, el principio de adyacencia predice que ahí se producirá un efecto de inducción cuyas magnitud y dirección quedarán determinadas más por el objeto-inducción que esté a la misma distancia aparente que el objeto-prueba y menos por el que esté separado de él; cuando el objeto-prueba esté en medio, a igual distancia de los dos objetos-inducción opuestos, o no habrá efecto o será mínimo.

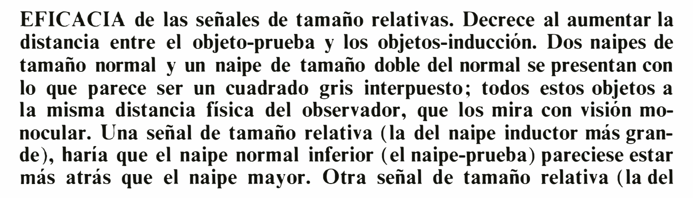
En uno de estos experimentos sobre la adyacencia en profundidad, manipulábamos los elementos de la ilusión de Ponzo, que es una ilusión óptica en la que dos gruesos trazos o palotes se inscriben en un ángulo o cuña: aunque los dos palotes son de idéntico tamaño, el más próximo al vértice parece mayor que el otro. Nosotros presentamos dos ángulos (los objetos-inducción) simultáneamente, a diferentes distancias, uno con el vértice hacia la derecha y otro con él hacia la izquierda, y presentamos los palotes (el objeto-prueba) a una de tres distancias: a la del ángulo más alejado, a la del ángulo más próximo y a distancia media entre ellos *(véase A en la ilustración de la página opuesta).* Tal como lo predecía el principio de adyacencia, la dirección de la ilusión se iba invirtiendo a medida que los palotes eran traídos del plano más apartado (donde había un promedio de ilusión del 10 por ciento en una dirección) al más próximo (donde se producía un promedio de ilusión del 10 por ciento en la otra dirección).

Un experimento realizado por Robert E. Newton y por mí mismo, basado en la ilusión de la barra y el marco, dio resultados parecidos. Aquí una barra vertical rodeada de un marco inclinado parece estar ella misma inclinada en sentido opuesto al de la inclinación física del marco *(B en la ilustración).* Una vez más, la ilusión se va invirtiendo según como se coloque, en profundidad, la barra vertical (el objeto-prueba). La inclinación promedio indicada por nuestros sujetos fue de cuatro grados en sentido contrario al de las agujas del reloj para la posición más alejada y de dos grados en el sentido de las agujas del reloj para la posición más próxima. Aunque las diferencias en el efecto de inducción entre la posición más distante y la menos no parezcan ser grandes para ninguna de estas ilusiones, resultan sustanciales si se las evalúa comparándolas con la magnitud normal de tales ilusiones.

En otros dos experimentos sobre la adyacencia en profundidad medimos el movimiento inducido en un punto-prueba por objetos-inducción que producían efectos opuestos. Con Michael Koslow investigué la magnitud y la dirección del movimiento inducido en un punto estacionario por dos marcos que oscilaban horizontalmente, en fases opuestas, a dos veces la distancia del observador *(C en la ilustración de la página opuesta).* Nuevamente hallamos que la dirección del movimiento inducido del punto era determinada por el marco inductor a la misma distancia que el punto, y que la magnitud de la inducción se reducía cuando el objeto-prueba se hallaba en el medio, a igual distancia de los dos marcos inductores. Jerome D. Tietz y yo mismo comprobamos la moción inducida, en un punto-prueba que se movía verticalmente, por objetos-inducción consistentes en dos puntos que oscilaban en sentido horizontal *(D en la ilustración).* El conjunto de los puntos-inducción más lejanos se movía hacia la derecha cuando el punto-prueba se movía hacia arriba, y el conjunto más cercano se movía hacia la izquierda cuando el punto-prueba se movía hacia arriba. Los resultados estaban de acuerdo con el principio de adyacencia: el punto-prueba parecía oscilar entre la parte superior izquierda y la parte inferior derecha cuando estaba en el plano más lejano, y entre la superior derecha y la inferior izquierda cuando estaba en el plano más próximo. El efecto de adyacencia era amplio: había unos 35 grados de diferencia entre las dos direcciones percibidas del movimiento del punto-prueba.



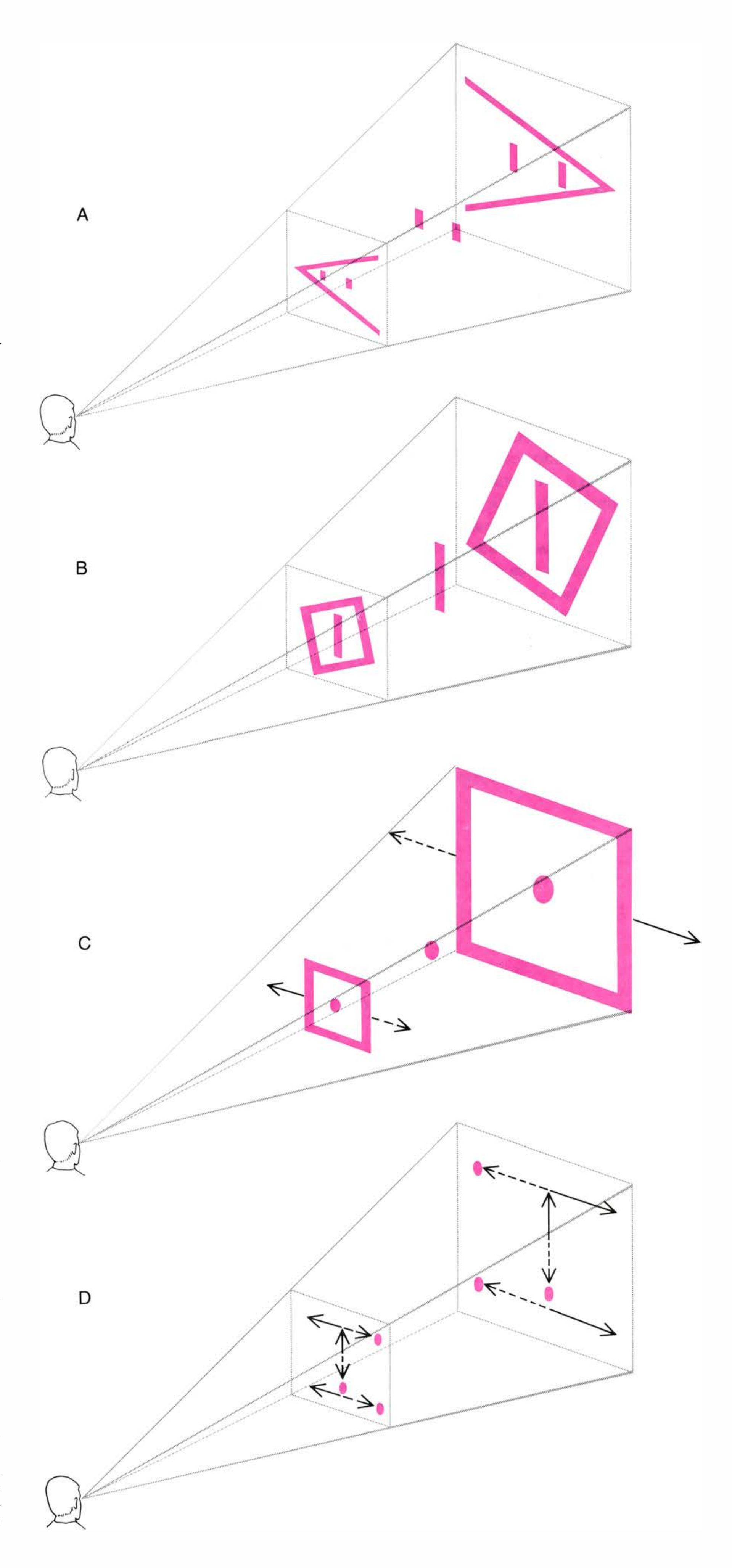




EFICACIA de las señales de tamaño relativas. Decrece al aumentar la distancia entre el objeto-prueba y los objetos-inducción. Dos naipes de tamaño normal y un naipe de tamaño doble del normal se presentan con lo que parece ser un cuadrado gris interpuesto; todos estos objetos a la misma distancia física del observador, que los mira con visión monocular. Una señal de tamaño relativa (la del naipe inductor más grande), haría que el naipe normal inferior (el naipe-prueba) pareciese estar más atrás que el naipe mayor. Otra señal de tamaño relativa (la del naipe inductor normal situado arriba) haría que el naipe-prueba pareciese estar a igual distancia que el naipe normal de arriba. Estas dos señales de tamaño relativas se han hecho incompatibles por la señal interpositiva del cuadrado gris. El naipe grande influía más cuando el naipe-prueba estaba cerca de él *(a),* es decir, el naipe-prueba parecía estar más allá, por detrás de los dos naipes inductores. Estos influían casi lo mismo cuando el naipe-prueba se hallaba a igual distancia de los dos *(b);* entonces, el naipe-prueba parecía estar entre los dos naipes-inducción.

En estos cuatro experimentos sobre la adyacencia en profundidad el efecto del objeto-inducción adyacente fue mayor que el del más distante, pero en todos los casos el objeto-inducción más distante siguió ejerciendo alguna influencia. La prueba pone de manifiesto que la disminución de la eficacia de una señal visual, al irse aumentando la separación, es gradual: una función continua mejor que una función discreta.

Nosotros modificamos tres de los experimentos sobre adyacencia en profundidad (la ilusión de Ponzo, la ilusión de la barra y el marco y el punto dentro de un marco no fijo) presentando sólo un objeto-inducción cada vez. Con el objeto-inducción lejano, se presentaba el objeto-prueba en el mismo plano, o a una o a dos veces la distancia por delante de él; cuando se presentaba el objeto-inducción cercano, el objeto-prueba estaba o en el mismo plano, o alejado del mismo a una o a dos distancias. En los tres casos, el efecto de inducción disminuía rápidamente al ir aumentando la distancia por delante del objeto-inducción, pero no cuando aumentaba la distancia por detrás del mismo. Posiblemente la falta de una correspondiente disminución para las posiciones de detrás del objeto-inducción se deba a que se dé una relación inversa entre la eficacia de señales absolutas en torno al objeto-prueba y la distancia que separa a éste del observador. Según esta hipótesis, a medida que el objeto-prueba se mueve hacia el observador (por delante del objeto-inducción) tanto la disminuida efectividad de la inducción predicha por el principio de adyacencia como la aumentada efectividad de las señales absolutas contribuyen a eliminar el efecto de inducción. Confor-



EL PRINCIPIO SE COMPRUEBA presentando binocularmente, a la vez y a diferentes distancias, dos objetos inductores que produzcan efectos opuestos en un objeto-prueba colocado en el plano de uno de los inductores o entre estos dos. (Sólo se presenta cada vez un objeto-prueba.) El efecto del ángulo *(A)* consiste en hacer que la barra vertical más próxima al vértice parezca más larga que la otra. El efecto del marco inclinado *(B)* es hacer que la barra vertical parezca inclinada. El marco, al moverse *(C)*, hace que el punto estacionario parezca moverse horizontalmente. Los puntos inductores que se mueven en sentido horizontal *(D)* hacen que el punto que se mueve en sentido vertical parezca moverse en diagonal. En cada caso, los dos objetos-inducción son de tamaño tal que subtienden el mismo ángulo visual y, por tanto, resultan de igual tamaño en la retina. Con todo, los dos objetos-inducción orientados en sentidos opuestos no se anulan. De acuerdo con el principio de adyacencia, el objeto inductor situado a la misma profundidad aparente que el objeto-prueba determina la dirección de las dos ilusiones estáticas *(A, B)* de los movimientos aparentes *(C, D)*

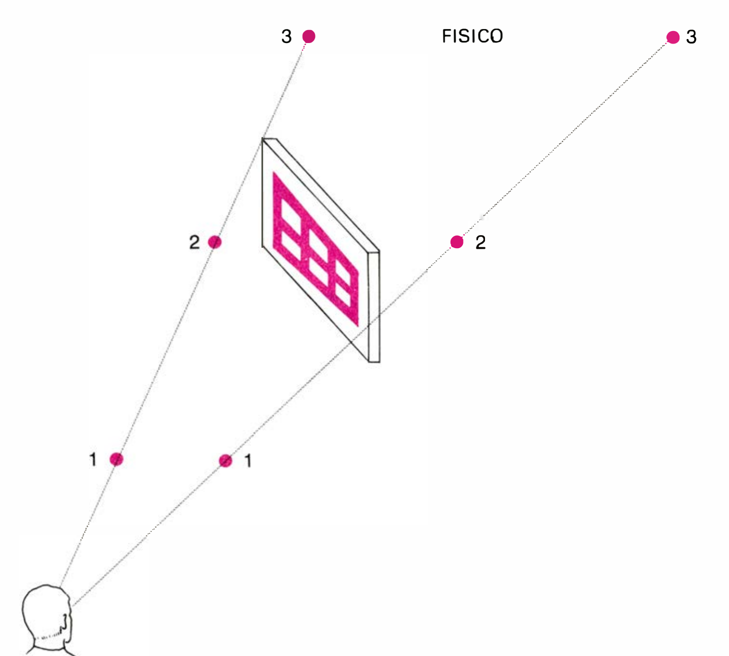
me el objeto-prueba se va apartando del observador, mengua probablemente la eficiencia de las señales relativas (inducción) y de las señales absolutas, de suerte que el efecto de inducción sigue siendo casi el mismo.

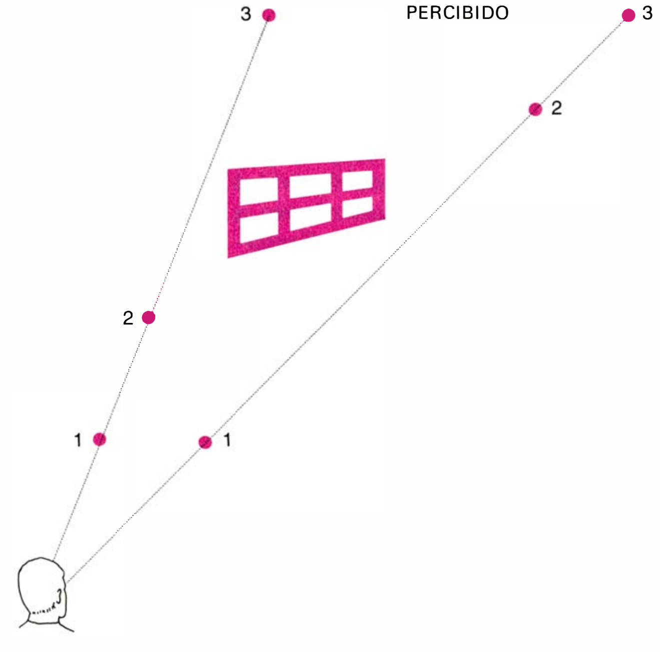
Hallamos, sin embargo, que desplazando un objeto-prueba detrás del objeto-inducción se reduce de hecho la magnitud de un efecto particular de inducción: el producido sobre los objetos de alrededor por la observación binocular de una ventana trapezoidal de Ames presentada con un sesgo ilusorio. El psicólogo Adelbert Ames, junior, ideó y compuso un marco de ventana trapezoidal que, presentado con un determinado sesgo en profundidad, tiende a parecer rectangular e inclinado en la dirección opuesta. Nosotros presentamos esta ventana bajo iluminación indirecta, con su extremo menor unas veces hacia la derecha y otras hacia la izquierda, pero siempre físicamente más cerca del observador que el extremo más ancho; pues bien, el extremo físicamente más próximo parecía estar más distante que el otro, aun cuando se miraba la ventana con visión binocular. Presentamos, a la vez que la ventana, dos puntos luminosos, uno a la derecha de la ventana y otro a la izquierda de ella *(véase la ilustración inferior).*

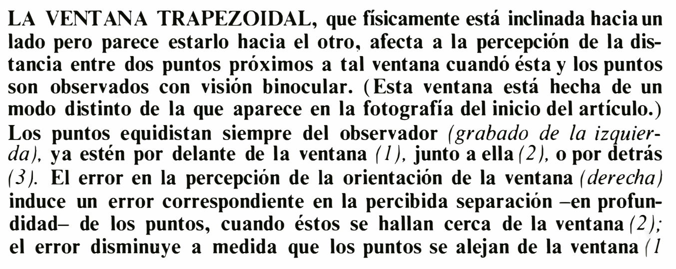
Este par de puntos luminosos fue presentado a siete diferentes distancias del observador (sólo tres de ellas se muestran en la ilustración). Aunque se movían las dos luces al mismo tiempo y se las mantenía siempre equidistantes del observador, el error de éste en el percibir la orientación de la ventana indujo un error sustancial en su percepción de las distancias relativas de los dos puntos.

Cuando ambos estaban aproximadamente a igual distancia que el centro de la ventana (a 305 centímetros del observador), la media de profundidad notificada entre el punto de la izquierda y el de la derecha fue de unos 60 centímetros. Alejando cada vez más los puntos por delante de la ventana o por detrás de ella, disminuía notoriamente este error en la percepción de la distancia, hasta hacerse muy pequeño (y parecer que los puntos estaban casi equidistantes) cuando se los ponía físicamente a unos 75 centímetros por delante o por detrás de la ventana. O sea, que sólo cuando los puntos luminosos estaban aproximadamente a la misma distancia que la ventana, repercutía el error mayor –el de percibir engañosamente la orientación de la ventana– en un error al percibir la separación de los dos puntos en profundidad. Quedaba así palmariamente demostrada, para la señal relativa de disparidad binocular, la tendencia del sistema visual a dar más importancia a las señales situadas entre objetos adyacentes que a las señales situadas entre objetos más separados.

¿Cómo podría darse razón de los efectos de adyacencia? Los efectos en un plano frontal paralelo podrían explicarse por la mengua de agudeza visual que se produce al irse separando la imagen del centro de la retina o *fovea centralis,* que es la parte más densamente poblada de células receptoras sensibles a la luz. A las señales entre objetos próximos el uno al otro en el campo visual podría darles más importancia el sistema visual que a las señales entre objetos separados, porque aquéllas tienen mayor claridad o son representadas con mayor precisión en el sistema nervioso. Los efectos de adyacencia en profundidad tendrían una explicación similar. Como los objetos que están fijos son vistos con particular claridad a resultas de la acomodación y convergencia de los ojos, el sistema visual puede que confíe más en las señales entre objetos que están a distancia fija que en las que están entre objetos a otras distancias. Una explicación diferente, pero algo relacionada con ésta, podría ser la de que los efectos de adyacencia dependen de la atención. A las señales entre objetos adyacentes se les daría más importancia que a las señales entre ob-






LA VENTANA TRAPEZOIDAL, que físicamente está inclinada hacia un lado pero parece estarlo hacia el otro, afecta a la percepción de la distancia entre dos puntos próximos a tal ventana cuando ésta y los puntos son observados con visión binocular. (Esta ventana está hecha de un modo distinto de la que aparece en la fotografía del inicio del artículo.) Los puntos equidistan siempre del observador ***(grabado de la izquierda),*** **ya estén por delante de la ventana** ***(1),*** **junto a ella** ***(2),*** **o por detrás** ***(3).*** **El error en la percepción de la orientación de la ventana** ***(derecha)*** **induce un error correspondiente en la percibida separación –en profundidad– de los puntos, cuando éstos se hallan cerca de la ventana** ***(2);*** **el error disminuye a medida que los puntos se alejan de la ventana** ***(1 o 3).*** **Los resultados concuerdan con el principio de adyacencia. Para un punto cercano, en profundidad, a la ventana** ***(2)*** **la señal de disparidad binocular entre él y la parte adyacente de la ventana –no la parte desplazada– determina la distancia aparente del punto: el punto de la derecha se halla físicamente más allá, y parece también estarlo, que el extremo derecho de la ventana, y el punto de la izquierda está físicamente delante o más acá del extremo izquierdo, y así parece estarlo. Cuando los puntos se desplazan en profundidad apartándose de la ventana, las señales binoculares directamente entre los puntos (mejor que entre cada punto y la ventana) cobran mayor importancia, influyen más y los puntos tienden a ser percibidos correctamente como equidistantes.**

jetos separados porque es más fácil atender simultáneamente a objetos que están cerca unos de otros.

Examinamos la explicación por la atención mediante unos experimentos con naipes. En estos experimentos, provocábamos juicios sobre la separación que pareciese haber, en profundidad, entre el observador y cada uno de los objetos que le presentábamos. Partíamos de la hipótesis de que, si la atención era importante, pidiendo al sujeto que juzgara la posición del naipe-prueba respecto a un determinado naipe-inducción, aumentaría la atención prestada a la señal de tamaño relativa situada entre el naipe-prueba y aquel naipe-inducción. La aumentada atención debería potenciar, a su vez, la eficacia de la señal relativa situada entre los dos naipes y reflejarse, por tanto, en la aparente distancia del naipe-prueba. Analizamos los resultados para verificar esta hipótesis y hallamos que el "ponerse a la tarea", una forma de atención, daba razón del 18 por ciento del efecto de adyacencia en el primer experimento *("a" en la ilustración de la página 70)* y del 35 por ciento del efecto en el segundo experimento *(b)*.

Una prueba más directa de la explicación por la atención la llevamos a cabo mediante el procedimiento de los puntos oscilantes vertical y horizontalmente en una disposición tridimensional *(D en la ilustración de la página 71)*. Pedimos a los observadores que indicaran la dirección del movimiento aparente del punto-prueba mientras atendían a un par de puntos inductores e ignoraban el otro par. Medimos la capacidad de atención para dar razón de los efectos de adyacencia, anotando cuánto cambiaba la dirección aparente según se dirigiese la atención a uno de los pares de puntos-inducción o al otro. La atención explicaba alrededor de la mitad del total del efecto de adyacencia. (En conjunto, el efecto de adyacencia fue grande en este experimento concreto, y el cambio atribuible a la atención fue también notable.)

Bajo algunas condiciones, pues, la atención es la causante de una parte sustancial –pero no del total– del efecto de adyacencia. El que no pueda explicarse por ella la totalidad del efecto sugiere que la adyacencia y la atención se basan en procesos diferentes. En cuanto a la claridad perceptual, acaso contribuya a los efectos de adyacencia, pero no puede explicar su magnitud. Por ejemplo, efectos de adyacencia en profundidad se presentan (en situaciones como la del experimento del punto oscilante) aun cuando la separación entre los objetos inductores es menor que la profundidad del foco del ojo, de suerte que ambos objetos inductores se ven claramente. Los experimentos abonan la conclusión de que, si bien factores como la atención voluntaria y la claridad visual pueden contribuir a que se produzcan fenómenos de adyacencia, queda una porción esencial de efectos de adyacencia que requiere que se postule alguna clase de proceso de valoración inconsciente determinada por la percibida separación espacial de los objetos.

El papel de la adyacencia espacial, en el proceso por el que la información proveniente de fuentes diversas se integra en el sistema visual, es quizás el mismo tanto si la información es complementaria como si es opuesta. No debería, sin embargo, infraestimarse la frecuencia con que las señales ópticas entran en conflicto entre sí. Algunos de los conflictos entre señales de nuestra investigación fueron poco comunes, tales como los producidos por la señal interpositiva y por la ventana trapezoidal. Pero otros tipos de conflictos entre señales resultan inevitables en la experiencia visual de cada día; de ellos, los hay entre señales relativas y señales absolutas, en general, y los hay entre diferentes señales relativas de movimiento, en particular. Por ejemplo, siempre que dos objetos situados a la misma distancia aparente se mueven uno respecto al otro, hay una señal de movimiento relativa, que está en conflicto con una por lo menos de las señales de movimiento absolutas. Siempre que tres objetos que estén a la misma distancia aparente se muevan diferentemente cada uno en relación a los demás, entrarán en conflicto varias señales de movimiento relativas. La integración de la información conflictiva parece ser un requisito normal para el sistema visual.

Juntas, las señales absolutas y las relativas constituyen toda la información de estímulo que se le ofrece al sistema visual. Las señales absolutas sostienen la percepción de las características de un objeto independientemente de las características de otros objetos; las señales relativas reflejan las interrelaciones perceptivas entre objetos. En ausencia de señales relativas, el mundo perceptivo constaría de partes independientes y sería muy difícil responder con sentido a tal mundo fragmentado. Para apreciar el grado de dificultad, considérese el caso, discutido por Johansson, de un disco rodante cuyas únicas partes visibles son varios puntos de luz en el borde y otro en el centro *(véase la ilustración de la página siguiente)*. Las trayectorias de los movimientos absolutos de los puntos serían una línea recta (para el punto central)

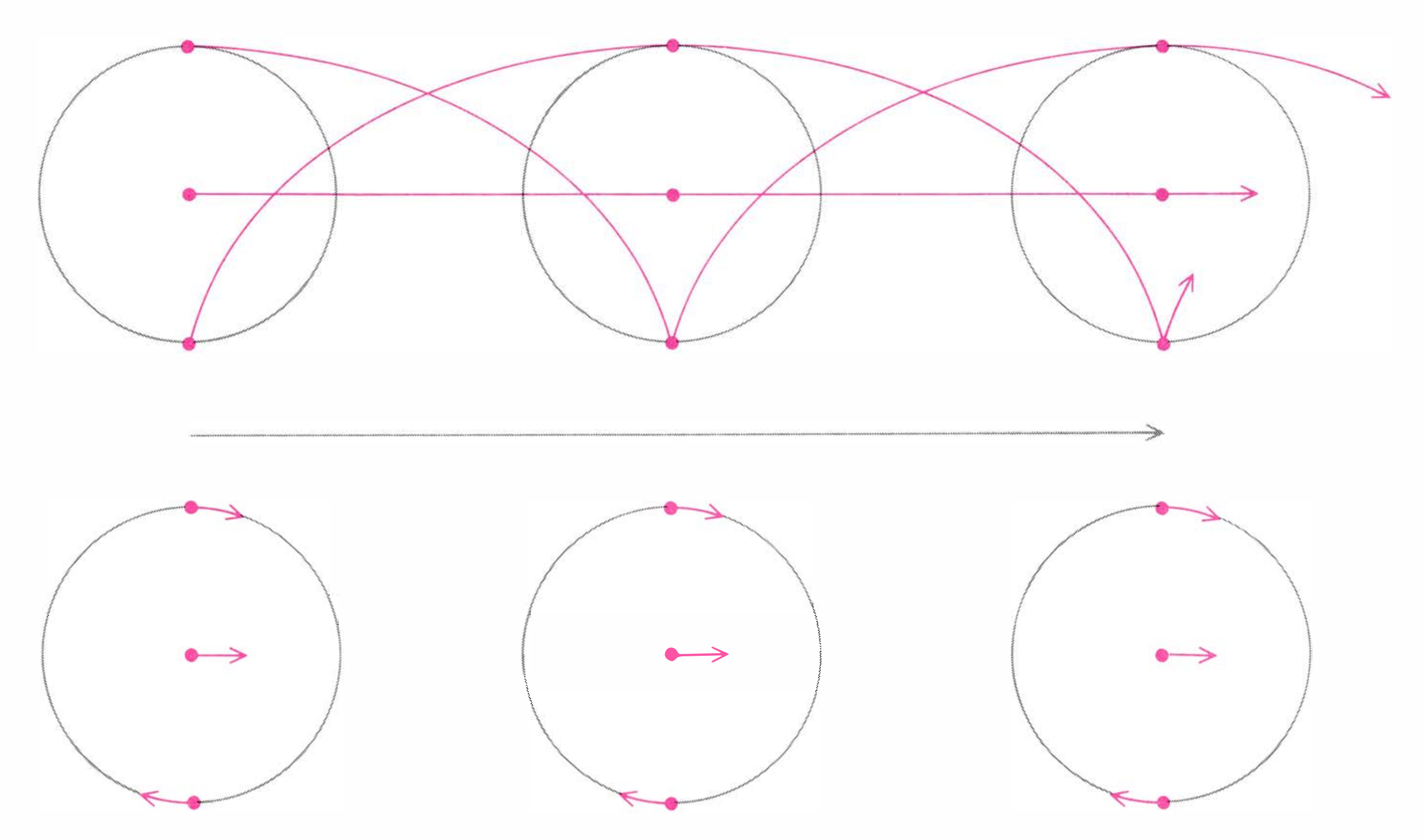
y curvas cicloides (para los puntos del borde). Una percepción basada en estos movimientos absolutos sería compleja, y se haría más compleja aún si se aumentase el número de puntos visibles en el borde. Pero, de hecho, los movimientos relativos de los puntos del borde y el punto central producen la percepción de un objeto unitario (un disco) que se mueve en una dirección (hacia la derecha) definida por sus vectores de movimiento común.

Generalizando a partir de este ejemplo, puede verse que se deriva una gran sencillez perceptiva del proceso organizador que agrupa estímulos reuniéndolos en objetos perceptivos, determinados por la posición relativa y el movimiento relativo y quizá también por otras señales relativas. Es mucho más sencillo percibir un objeto que se mueve como una unidad, que percibir un gran número de movimientos de puntos diferentes con independencia perceptiva.

¿En qué contribuye la adyacencia espacial a esta especie de organización perceptiva? La tendencia a percibir el movimiento relativo, y no el movimiento absoluto, de los puntos espacialmente adyacentes, y así a percibir estímulos organizados (tales como un disco giratorio), concuerda con el principio de adyacencia. Según lo hicieron notar insistentemente los psicólogos de la Gestalt, la proximidad en un plano frontal paralelo es un importante factor en la agrupación perceptiva. De acuerdo con el principio de adyacencia, la proximidad es un factor importante no sólo en el espacio de dos dimensiones sino también en el tridimensional. El grado en que las señales relativas de posición, movimiento y quizás otras características puede organizar estímulos reuniéndolos en grupos perceptivos, u objetos, depende de la proporción en que la eficacia de estas señales relativas disminuye al ir aumentando la separación espacial, es decir, depende de lo que podría denominarse la función de adyacencia. El mundo perceptivo tendería a fragmentarse demasiado si la curva de la función de adyacencia subiese muy bruscamente, y tendería a ser demasiado indiferenciado si la curva de esa función se mantuviese más o menos constante.

Muestras de estímulos no fijos, espacialmente separadas (dos discos rotatorios, por ejemplo), llegan a agruparse formando objetos perceptivamente distintos, en vez de constituir un confuso objeto complejo, gracias, por lo menos en parte, a efectos de adyacencia. Los movimientos comunes de las partes de cada objeto perceptivamente separado se convierten en los percibidos movimientos absolutos del objeto como un todo. Si los movimientos comunes de los distintos objetos percibidos son diferentes, esos movimientos pueden convertirse en movimientos relativos para más altos órdenes de organización visual. A resultas de este proceso, el mundo visual se va articulando y estructurando en vez de ser un caleidoscopio de piezas perceptivamente independientes. Es de esperar que el principio de adyacencia rija para esta especie de organización más global, así como rige para el más simple agrupamiento implicado en la percepción de un solo objeto.

Finalmente, puede ser que el principio de adyacencia esté en la raíz de algunas diferencias individuales en cuanto a la capacidad o al estilo del percibir. Investigando sobre los efectos de adyacencia hemos notado a menudo grandes diferencias entre los distintos individuos. Si se averigua que estas diferencias en las funciones de adyacencia individuales son estables durante un período de tiempo y son características de los distintos observadores podrá explicarse por ellas algo del por qué varía el modo en que los individuos perciben visualmente los objetos y perciben su mundo como compuesto de partes organizadas, e interrelacionadas.



SEÑALES RELATIVAS, que contribuyen a que adquieran sentido lo que, de lo contrario, serían estímulos visuales fragmentarios. Si la percepción de un disco rodante, invisible salvo en puntos luminosos de su periferia y de su centro, estuviera regida tan sólo por señales de movimiento absoluto, únicamente se verían la recta y las curvas trazadas por esos puntos luminosos *(arriba)*. La respuesta del sistema visual a los movimientos relativos y comunes de los puntos luminosos logra una percepción más simple, a saber: un objeto que rueda hacia la derecha *(abajo)*.



Uniones intercelulares

En algunos tejidos animales fundamentales la zona intercelular no consiste en un simple contacto entre membranas. Las células están unidas por estructuras especiales cuya arquitectura revela la microscopía electrónica

L. Andrew Staehelin y Barbara E. Hull

Los tejidos de los organismos superiores están constituidos por un gran número de células que han renunciado en parte a su independencia para funcionar integradamente como un conjunto unificado y coherente. Con el fin de llevar a cabo esta unidad y para facilitar las interacciones necesarias entre las células, ciertas regiones de la superficie celular se han especializado con objeto de establecer el contacto intercelular. Estas especializaciones superficiales, denominadas uniones intercelulares, son esenciales para el desarrollo y el funcionamiento normal de todas las formas de vida superiores.

Si le encomendaran a uno la tarea de construir un organismo a partir de células individuales, podría prever la necesidad de tres tipos funcionales de uniones intercelulares: (1) uniones impermeables que permitieran a un organismo o a un órgano mantener un medio interno químicamente diferente del medio circundante, (2) uniones adherentes que facilitaran la adherencia entre las células de un tejido, reforzando así su integridad física, y (3) uniones de comunicación que hicieran posible el intercambio de nutrientes y moléculas señalizadoras, pudiendo así coordinar las actividades de las células. En los veinte últimos años, los biólogos celulares no sólo han confirmado la existencia de estos tres tipos de uniones sino que también han acumulado una gran riqueza de información acerca de su arquitectura molecular.

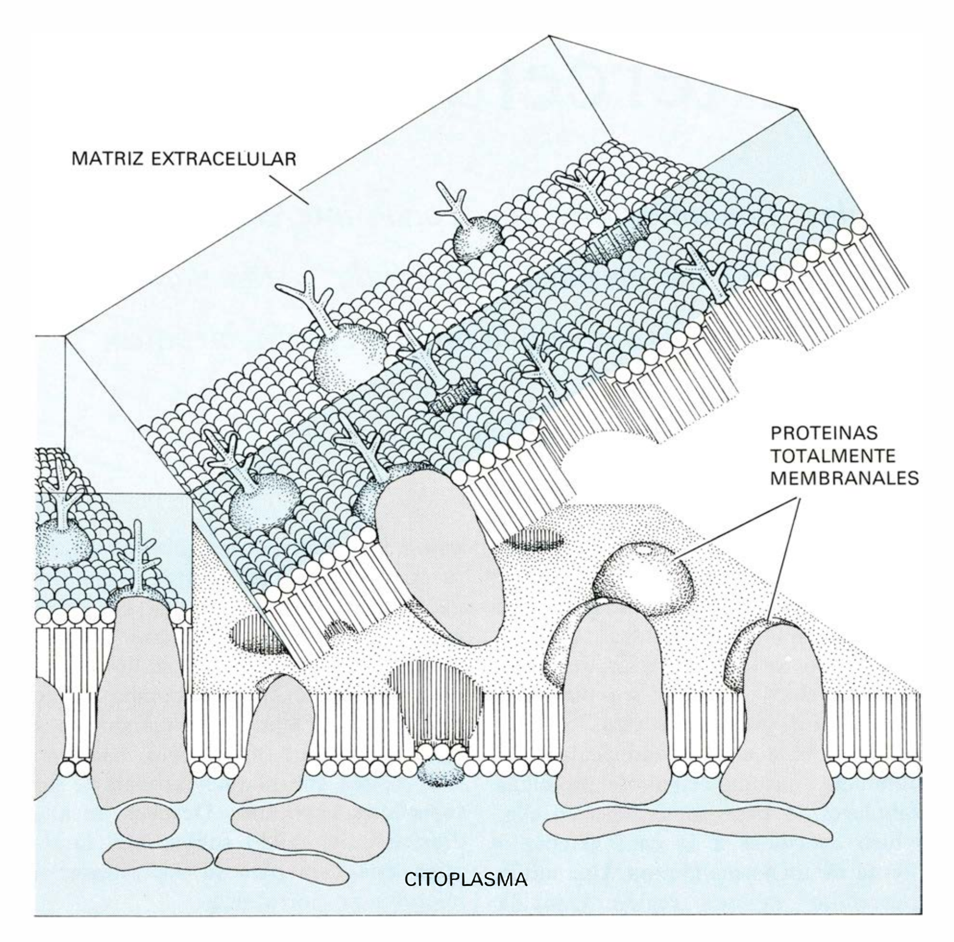
Todas las células vivas están limitadas por una membrana plasmática, una envoltura delgada que regula de manera selectiva el flujo de nutrientes y de iones que la atraviesa, del interior de la célula al medio externo o en dirección opuesta. La membrana está formada por una doble capa de moléculas fosfolipídicas con moléculas de proteínas globulares, o bien sumergidas en ella, o bien asociadas a la cara externa o interna de dicha membrana. Una unión intercelular es una región reducida de la membrana plasmática que está especializada para funcionar como uno de los tres tipos de uniones anteriormente mencionados. Dado el pequeño tamaño de las células y, por tanto, de las uniones, resulta necesario el empleo de microscopio electrónico para observar la estructura fina de dichas uniones.

Fundamentalmente, existen dos métodos de preparación del espécimen (u objeto de estudio) para adecuar las células a su ulterior observación al microscopio electrónico: la ultramicrotomía y la congelación-fractura. Para realizar cortes ultramicrotómicos, primero se ha de estabilizar químicamente las células, teñirlas luego, deshidratarlas, incluirlas en plástico y, finalmente, cortarlas en rodajas sumamente finas con un ultramicrotomo. La observación al microscopio electrónico de los cortes en los que aparece una región de unión permite determinar algunos aspectos de la organización tridimensional de la unión. La técnica de congelación-fractura implica la congelación de la muestra de tejido a menos 150 grados Celsius en freon líquido y su posterior transferencia a una cámara que se lleva a un gran vacío mediante una bomba. Luego se fractura el tejido atravesándolo con una cuchilla de microtomo, haciendo una réplica con platino-carbono de sus superficies expuestas. Después de una digestión del tejido subyacente, la réplica está lista para su observación al microscopio electrónico.

La congelación-fractura ha aportado nuevas dimensiones al entendimiento de la estructura de la membrana, ya que, a menos 100 grados C, el proceso de fractura parte las membranas celulares a lo largo del plano central de la bicapa lipídica. La mitad de la membrana se queda unida al citoplasma subyacente y la otra mitad se queda unida a la matriz extracelular. Las moléculas de proteína globular sumergidas en la membrana (proteínas totalmente membranales) se visualizan como bultitos que sobresalen de la superficie lisa de una u otra de las caras de semimembrana (lo que nos revela que las proteínas se extienden a lo largo del plano central de la bicapa). Si nos hallamos en el caso de una región de unión intercelular, las proteínas totalmente membranales presentan configuraciones específicas, que se pueden observar en las réplicas de congelación-fractura. En definitiva, la congelación-fractura y la ultramicrotomía pueden aportar información complementaria acerca de los elementos estructurales de estas regiones tan especializadas de la superficie celular.

Las uniones intercelulares desempeñan un papel de particular importancia en los tejidos epiteliales que recubren la superficie del cuerpo y revisten sus cavidades. Por ejemplo, las células epiteliales del intestino delgado que revisten la cavidad interna del intestino (la

REGION DE CONTACTO INTIMO entre dos células, pertenecientes al revestimiento del intestino delgado de rana, aumentada 135.000 diámetros en la electromicrografía de la página opuesta. El espécimen se preparó por el método de congelación-fractura, en el que las células se quiebran a temperaturas sumamente bajas y luego se realiza un sombreado con platino y carbono para poner de manifiesto su detalle tridimensional. Aquí el proceso de fractura ha eliminado la célula superior en su mayoría, dejando tan sólo un pequeño fragmento de la membrana celular asociado a la superficie de las célula inferior. En la parte superior de la micrografía, se pueden observar microvellosidades, unas prolongaciones tubulares de la membrana de la célula inferior que incrementan muchísimo la superficie de absorción del intestino. En la zona inmediatamente inferior a la mencionada se observa una red de crestas, a modo de celdillas de panal, que se llama unión hermética, unión que forma un puente entre las membranas de las células adyacentes. La anchura del espacio intercelular viene indicada por el fuerte "acantilado" *(parte inferior izquierda),* **que representa un lugar donde el plano de fractura salta de la célula superior a la inferior.**



CONGELACION-FRACTURA, que parte la membrana celular a lo largo de un plano central que atraviesa su bicapa de moléculas lipídicas; este mecanismo sirve para poner de manifiesto su estructura interna. Como se indica en este diagrama esquemático, la mitad de la membrana permanece asociada con el citoplasma subyacente y la otra mitad queda unida a la matriz extracelular. Las moléculas de proteínas globular que normalmente se encuentran sumergidas en la bicapa se ven entonces puestas al descubierto y, en las réplicas de platino-carbono, aparecen como bultos repartidos por la superficie lisa de cada cara de las semimembranas. Si la región es la de una unión intercelular, estas proteínas adoptan configuraciones relacionadas con la arquitectura de la unión.

luz) absorben selectivamente los nutrientes y los transportan a través de su célula para que puedan penetrar en el torrente circulatorio. Está claro que este sistema de transporte sería francamente ineficaz si los nutrientes absorbidos por las células, con un gasto de energía metabólica por su parte, volvieran a filtrarse hacia la luz del intestino por los espacios intercelulares. Las uniones intercelulares denominadas uniones estrechas o herméticas ("tight junction") sirven para impedir el libre paso de moléculas a través de un epitelio.

En la región correspondiente a una unión estrecha, las membranas plasmáticas de las dos células epiteliales adyacentes parecen fusionarse, formando una región de íntimo contacto que constituye un aro que ciñe a cada célula. Como demostraron los trabajos de Marilyn G. Farquhar y George E. Palade, de la Universidad Rockefeller, cuando se expone la superficie interna de un epitelio a una solución de una sustancia marcadora densa a los electrones, como la proteína ferritina, el marcador penetra fácilmente en el espacio intercelular, pero se detiene a nivel de la unión estrecha. En imágenes de cortes ultramicrotómicos de uniones estrechas, se observa con claridad que las dos membranas plasmáticas no están unidas por un cemento sino que más bien se hallan fusionadas en una serie de puntos.

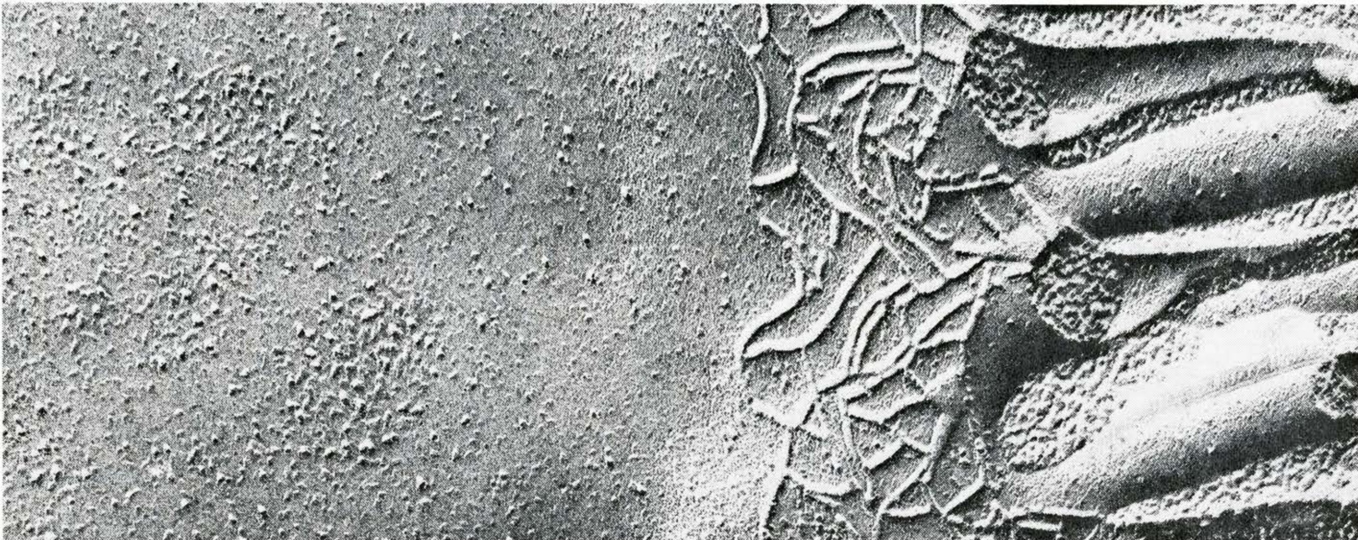
La organización estructural bidimensional de las membranas en la región de una unión estrecha se dedujo en 1968 a partir de estudios de congelación-fractura realizados por uno de los autores (Staehelin), que se hallaba trabajando en el Departamento de Investigación Científica e Industrial de Nueva Zelanda, en colaboración con Tapen M. Mukherjee, de la Universidad de Otago. En las réplicas de uniones estrechas, realizadas por congelación-fractura, estas uniones se caracterizaban por una red de crestas en la cara de la media membrana citoplasmática y otra de hoyos complementarios en la cara de la media membrana externa. En el lugar donde el plano de fractura salta de la membrana plasmática de la célula inferior a la de la célula superior, parece como si las crestas ocuparan toda la anchura del conjunto de ambas membranas.

Las pruebas de que disponemos sugieren que las crestas de la unión estrecha están compuestas por dos filas de partículas fuertemente unidas (que serían proteínas totalmente membranales), contribuyendo cada una de las membranas plasmáticas adyacentes con la aportación de una fila de estas partículas. Estas filas de partículas entran en contacto mutuo, frente a frente, como si se formara una cremallera modificada que mantiene tan próximas las dos membranas que el espacio intercelular se halla obliterado. En consecuencia, las partículas forman líneas de unión (hebras de cierre) que bloquean físicamente el paso de moléculas a través del epitelio.

El número de hebras de cierre que poseen las uniones herméticas varía de un tejido a otro. Puesto que las hebras de cierre constituyen una barrera de permeabilidad, cuanto mayor sea el número de hebras, más impermeable será la unión. Se puede obtener un indicio aproximado de la hermeticidad del cierre midiendo la resistencia transepitelial: es decir, la resistencia de un epitelio al movimiento de iones a través de él. Philippa Claude y Daniel A. Goodenough, de la Facultad de Medicina de Harvard, han comprobado que existe una buena correlación entre la resistencia transepitelial de un tejido y el número de hebras de cierre que hay en las uniones estrechas entre sus células componentes. En los tejidos en que las uniones herméticas o estrechas comprenden tan sólo una o dos hebras de cierre como ocurre, por ejemplo, en el túbulo proximal del riñón, los epitelios ofrecen poca resistencia al paso de iones. En cambio, en las uniones donde hay un número de hebras igual o superior a seis, como en la vejiga urinaria, la resistencia eléctrica es elevada y se permite la formación de fuertes gradientes de concentración a través del epitelio.

En los diversos tejidos, las uniones estrechas también varían en cuanto al grado en que sus hebras de cierre realizan interconexiones unas con otras formando una red. Parece que la importancia cuantitativa de estas interconexiones es lo que determina hasta qué punto la red podrá responder a la tensión, al igual que la extensibilidad de una tela depende de cuán apretadamente esté tejida. Un hecho concorde con esta hipótesis es nuestro descubrimiento de que, en las células que mantienen una forma relativamente constante, hay una red poligo-





ESTAS VISTAS COMPLEMENTARIAS de uniones intercelulares proceden de la aplicación a la muestra de las técnicas de ultramicrotomía y congelación-fractura, respectivamente. Estas dos microelectrografías representan la misma región de contacto íntimo entre las membranas de dos células adyacentes del epitelio (capa celular externa) del intestino delgado. En la micrografía de un corte ultramicrotómico (la superior), se pueden distinguir tres tipos distintos de uniones intercelulares. Muy a la derecha, cerca de la base de las microvellosidades, parece que las membranas plasmáticas adyacentes de dos células se fusionan, formando una unión hermética. En el centro de la micrografía, hallamos el desmosoma de banda, marcado por densos entretejidos de filamentos finos que se adhieren a las superficies internas de las membranas adyacentes. Muy a la izquierda de la micrografía, vemos una unión muy teñida llamada desmosoma puntiforme. En la imagen de congelación-fractura (la inferior), se ve claramente la red de la unión hermética a la derecha. A la izquierda de la unión hermética, hay una zona poco destacada que contiene unas cuantas partículas: es el desmosoma de banda. Muy a la izquierda, los desmosomas puntiformes aparecen como parches de partículas con formas irregulares. Los tres tipos de unión constituyen el denominado "complejo de unión". Aumento: x17.000.

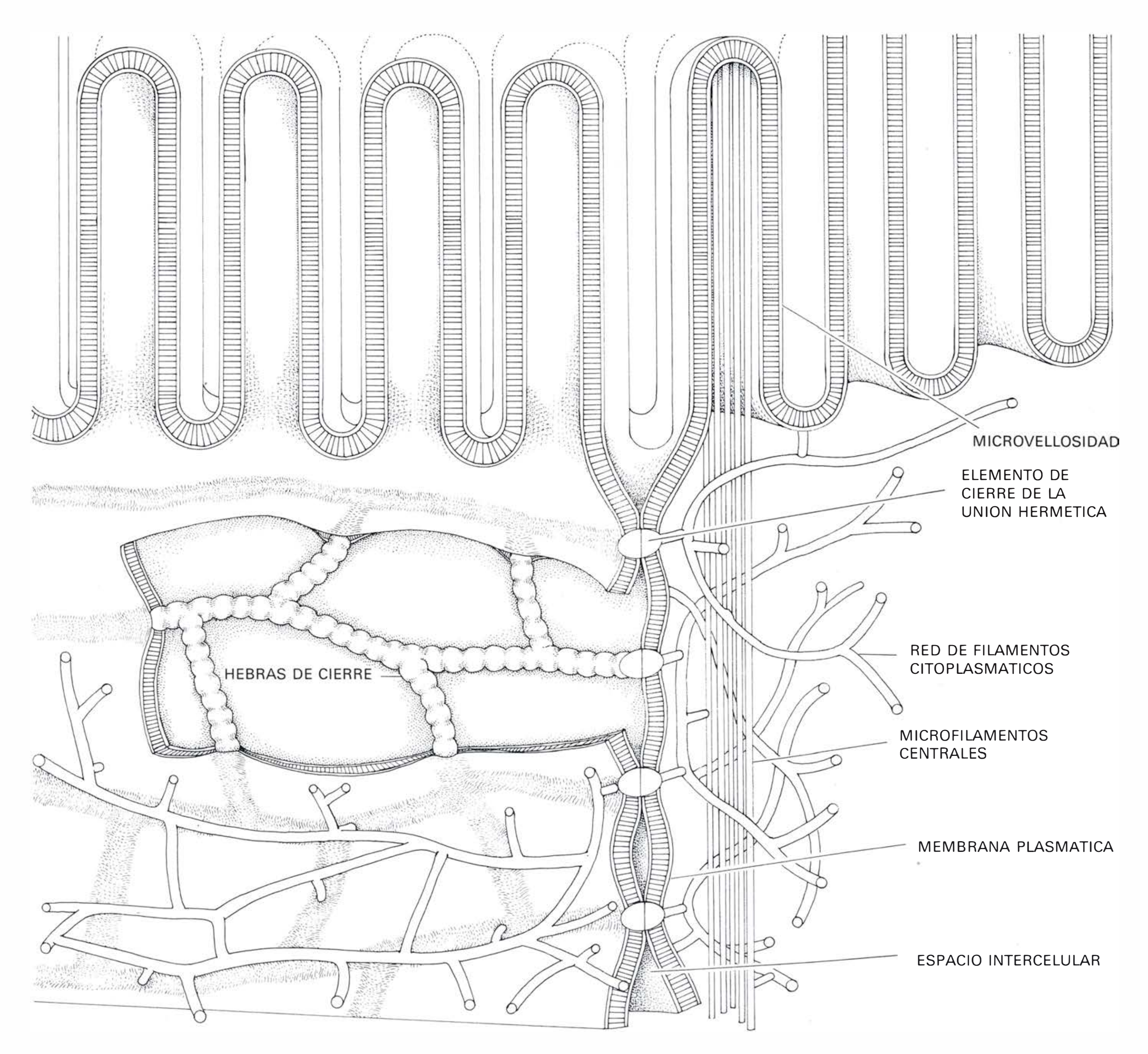


LA DETENCION DEL FLUJO de la sustancia opaca a los electrones, hidróxido de lantanio, a través del espacio intercelular se da a nivel de la unión hermética, como indica esta microelectrografía de células epiteliales de páncreas. Las uniones herméticas, que ejercen un bloqueo físico del paso de moléculas a través del espacio, revisten particular importancia en la prevención de la infiltración de iones y nutrientes a través del epitelio de un órgano. Aquí el elemento de cierre de la unión estrecha aparece como un punto de fusión entre las membranas contiguas. El aumento de la micrografía, proporcionada por D. S. Friend, de la Universidad de California, es de 127.000 diámetros.

nal de interconexiones regulares: así ocurre, por ejemplo, en las células epiteliales del intestino delgado. Por el contrario, las redes irregulares de hebras de cierre con sólo escasas interconexiones suelen pertenecer a células que se ven obligadas a expandirse para acumular un producto de secreción (como las células mucosas secretoras del estómago) o a estirarse periódicamente por efecto de una tensión (como las células epiteliales del intestino grueso).

El diseño de la unión estrecha resulta ser un ejemplo magnífico de ingeniería biológica a nivel celular. La organización de las hebras de cierre formando una red permite variar la hermeticidad del cierre según las necesidades fisiológicas del tejido. La red de cierre es increíblemente flexible y se puede someter a tensión, comprimir o torcer sin que se pierda, por ello, la capacidad de cierre. Por otra parte, posee un margen de seguridad inherente muy amplio, ya que la ruptura local de una o varias hebras de cierre de la unión estrecha tiene poco efecto sobre el conjunto de la red y, por tanto, sobre la hermeticidad del cierre.

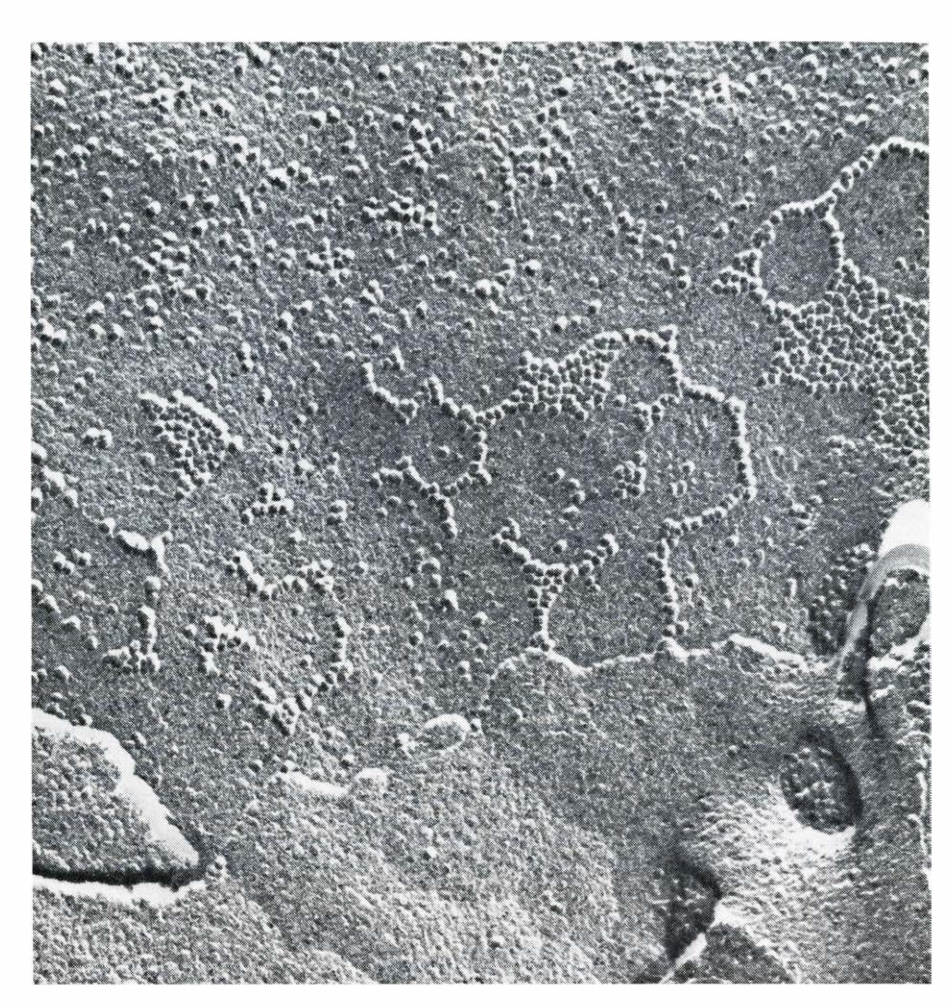
Como ya hemos comentado, la función principal de la unión estrecha es permitirle a un grupo de células mantener un medio interno diferente del medio que las rodea. Otro hecho, que no se suele ensalzar tanto, es que las uniones herméticas contribuyen significativamente a establecer la polaridad de la célula y de la membrana plasmática. Debido a que la bicapa lipídica de la membrana plasmática es semifluida a las temperaturas fisiológicas, muchas de las proteínas totalmente membranales tienden a migrar en el interior del plano de la membrana y a distribuirse al azar por toda la superficie celular. Sin embargo, la presencia de uniones herméticas introduce una barrera física a la migración de estas proteínas de la superficie externa de un epitelio hacia su superficie interna o viceversa. Si no existiera esta barrera, las proteínas que están implicadas en el



MODELO DE UNA UNION ESTRECHA, ilustrado en este diagrama esquemático. Las membranas de las células adyacentes se mantienen unidas por hebras de cierre que forman líneas de unión. Cada hebra de cierre está compuesta por dos filas de partículas muy próximas entre sí (cada célula aporta una fila) que se adhieren fuertemente a modo de una cremallera modificada. La ordenación de las hebras de cierre para formar una red concede a la unión una flexibilidad considerable y le permite mantener el cierre en diversas condiciones de tensión. Hay unos filamentos asociados a las superficies citoplasmáticas de las membranas que ponen en contacto las hebras de cierre y refuerzan la unión.

transporte de nutrientes hacia el interior de la células (procedentes de la luz intestinal) se mezclarían con las proteínas implicadas en el transporte de los mismos nutrientes hacia el exterior de la célula (hacia el torrente sanguíneo). Se han hallado uniones herméticas, incluso en etapas del desarrollo muy tempranas, por ejemplo en el estadio bicelular del embrión, y parece que estas uniones son fundamentales para un desarrollo embrionario normal.

Pero son varios los problemas intrigantes en torno a la estructura de las uniones herméticas que quedan aún por resolver. No sabemos cómo se ensamblan las subunidades proteicas para formar las redes en la membrana de las células adyacentes, proceso que precisaría una coordinación de la maquinaria de síntesis de ambas células. Tampoco sabemos qué factores provocan las variaciones del tipo de red en los diferentes tejidos. Las primeras pistas nos las han proporcionado unos estudios sobre el desarrollo de uniones herméticas llevados a cabo por Roberto Montesano, Alain Perrelet y Lelio Orci, de la Universidad de Ginebra, en colaboración con Daniel S. Friend, de la Universidad de California en San Francisco. Estos investigadores observaron la formación de la red típica en las células hepáticas de fetos de rata y comprobaron que sigue una secuencia ordenada de acontecimientos. En la región en que se está formando la unión hermética, se organizan unas cortas filas de partículas en torno a los bordes de unas depresiones de la membrana plasmática que tienen forma de celdillas de panal. Las filas de partículas se van interconectando progresivamente constituyendo una red laxa que poco a poco se va apretando, cada vez más, dando finalmente las crestas suaves, características de la unión hermética madura. Un análisis minucioso de estos acontecimientos a nivel molecular debería ser factible en los próximos años.



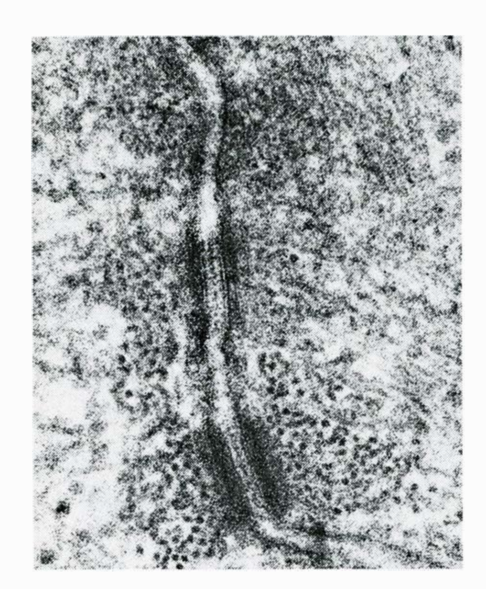
FORMACION DE UNA UNION ESTRECHA entre dos células hepáticas de feto de rata, en este espécimen preparado por congelación-fractura. La formación de la unión sigue una secuencia ordenada de acontecimientos. En primer lugar, grupos de partículas se organizan en filas que luego se ordenan en una red de hebras de cierre, a modo de celdillas de panal. Aquí observamos el inicio de formación de dicha red. El aumento correspondiente a la micrografía es de 102.000 diámetros.

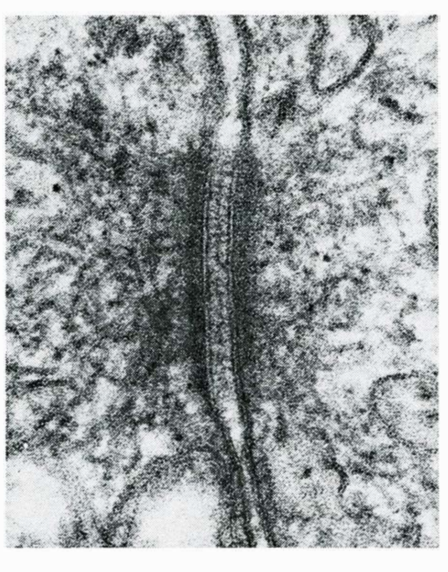
La adherencia de las células de un tejido se mantiene mediante uniones de conjunción denominadas desmosomas, los cuales permiten que grupos de células funcionen al unísono como una unidad estructural. Estas uniones se pueden clasificar en dos categorías principales: los desmosomas de banda y los desmosomas puntiformes, en razón de los tipos de filamentos citoplasmáticos con los que van asociadas.

El desmosoma de banda forma una franja que une a las células epiteliales adyacentes en la región inmediatamente inferior a la unión hermética. En esta zona de unión, el espacio intercelular está lleno de un material filamentoso fino. Por otra parte, a las superficies citoplasmáticas de ambas membranas plasmáticas adyacentes se encuentran asociados dos grupos de filamentos: un grupo en forma de manojo a lo largo de la cara interna de la membrana y otro que se origina en la unión, o en sus proximidades, y se extiende hacia el citoplasma con una configuración aplanada.

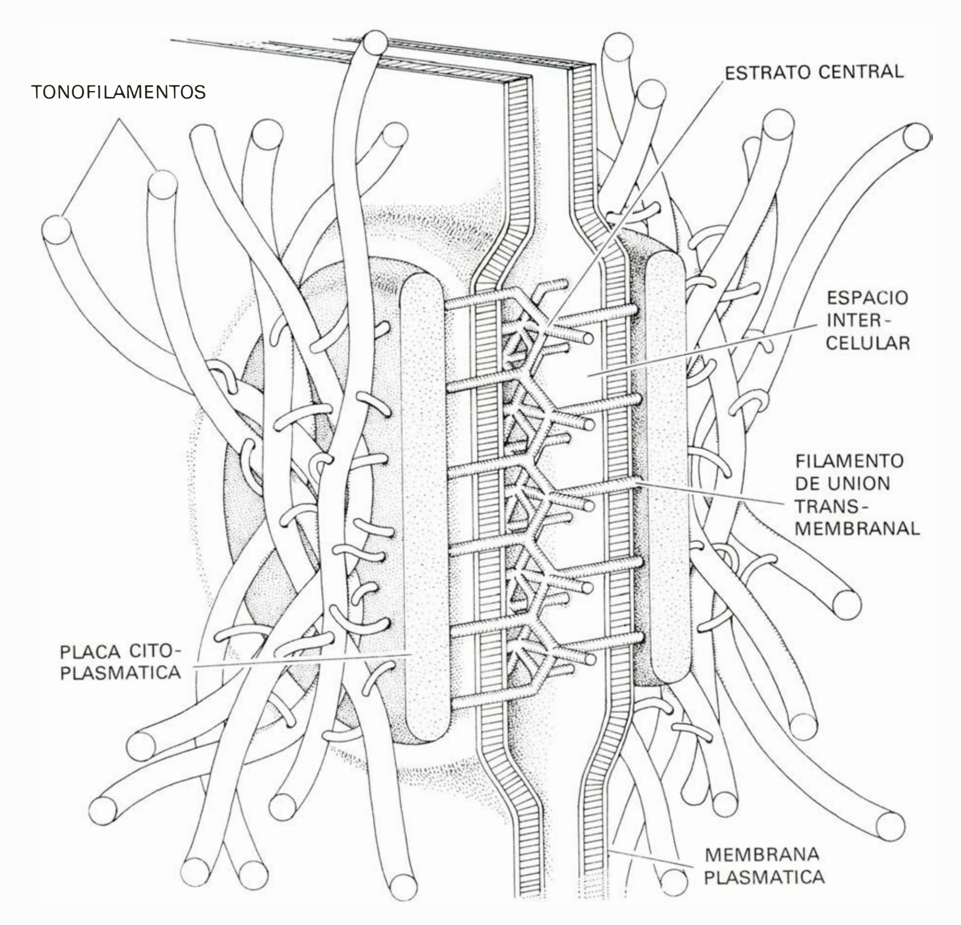
Los filamentos asociados con el desmosoma de banda tienen 70 angstroms de diámetro y contienen, según parece, actina, la principal proteína de las células musculares, lo que hace pensar que estos filamentos son capaces de contraerse. Efectivamente, Richard D. Rodewald y Morris J. Karnovsky, de la Facultad de Medicina de Harvard, y Mark S. Mooseker y Lewis G. Tilney, de la Universidad de Pennsylvania, demostraron que la región de la célula epitelial intestinal que comprende al desmosoma de banda se contrae en presencia de ATP (molécula de almacenaje de energía) y de iones calcio o magnesio. Debido a que las células epiteliales están físicamente unidas formando una capa coherente, no está claro cuál puede ser el papel de la contracción del desmosoma de banda que ciñe a cada célula. Al parecer, sirve para cerrar las lagunas formadas en el epitelio a consecuencia de la muerte de las células y de su desprendimiento. Los filamentos de 70 angstroms quizá sean también los responsables de los movimientos y de los cambios de forma de las capas epiteliales durante la constitución de los órganos en el embrión.

El segundo tipo de unión adherente es el desmosoma puntiforme, que fue descubierto en 1954 por Keith R. Porter, de la Universidad Rockefeller. Contrariamente a las uniones estrechas y a los desmosomas de banda, los desmosomas puntiformes no forman franjas o bandas alrededor de las células epiteliales sino que constituyen más bien puntos de contacto en forma de botón entre las membranas plasmáticas de las células adyacentes, de distintos niveles, análogos a unos roblones o puntos de soldadura. En el interior del desmosoma puntiforme, las dos membranas celulares adyacentes son estrictamente paralelas y están se-





SON DOS LOS TIPOS DE UNION ADHERENTE, el desmosoma de banda y el desmosoma puntiforme, responsables de los fuertes lazos mecánicos entre células, lazos que impiden su desgarre cuando el tejido se ve sometido a tensión. En la micrografía de la izquierda, aumentada 105.000 veces, el desmosoma de banda se halla en la parte superior y, debajo de él, vemos dos desmosomas puntiformes. Las fibras asociadas al desmosoma de banda, llamadas filamentos, de 70 angstroms, tienen un diámetro visiblemente inferior que los asociados al desmosoma puntiforme, llamados tonofilamentos. (Estos filamentos, cortados en sección transversal, aparecen como puntos negros.) La micrografía de la derecha, con un aumento de 135.000 diámetros, corresponde a un desmosoma puntiforme. En la región de la unión, revistiendo las superficies internas de las membranas celulares adyacentes, se hallan dos placas de tinción obscura, a las que se unen manojos de tonofilamentos. Hay otros filamentos más finos que se originan en las placas y se extienden a través de las membranas celulares para unirse a cierta estructura, llamada estrato central.



MODELO DE UN DESMOSOMA PUNTIFORME representado en grandes líneas en este diagrama. Los tonofilamentos, de 100 angstroms de diámetro, forman una red resistente a la tensión que se extiende por todo el interior de la célula. Estos tonofilamentos están unidos a las placas del desmosoma puntiforme a través de estructuras filamentosas vagamente definidas. Hay otros filamentos, llamados de unión transmembranal, que conecta las placas del desmosoma puntiforme a través del espacio intercelular. La unión sirve para acoplar las redes de tonofilamentos de las células adyacentes, permitiendo una disipación de las fuerzas divisorias por todo el tejido.

paradas por un espacio de aproximadamente 300 angstroms de anchura. El espacio está relleno de un material filamentoso: cuando se observan cortes ultramicrotómicos del mismo, parece estar zanjado en su parte media por una línea densa llamada estrato central. En la superficie citoplasmática de cada membrana plasmática hay una placa en forma de disco. A estas placas se hallan conectados unos filamentos de 100 angstroms de diámetro, llamados tonofilamentos. Los tonofilamentos no son contráctiles, pero, al parecer, forman un armazón estructural en el citoplasma celular que puede resistir bien la tensión. Hay manojos de tonofilamentos que se originan en la profundidad del citoplasma y que forman un bucle al atravesar las placas de los desmosomas puntiformes. Otros manojos se extienden paralelos a la membrana plasmática y prácticamente no varían su dirección al pasar por la placa.

Hay filamentos más finos que surgen de las propias placas y se prolongan desde la célula hacia el espacio intercelular, donde entran en conexión con el estrato central en una disposición alternada. Estos filamentos de "unión transmembranal" proporcionan un acoplamiento mecánico directo entre las redes de tonofilamentos de células epiteliales adyacentes, creándose de este modo una trama estructural continua para la totalidad del epitelio. Los estudios de congelación-fractura han confirmado la existencia de este acoplamiento mecánico entre células. Contrariamente a lo que ocurre con las membranas de los desmosomas de banda, que prácticamente no contienen partículas, las membranas del desmosoma puntiforme poseen placas discoidales, formadas por partículas fracturadas de manera irregular. Estas partículas se asemejan a unos filamentos que se hubieran roto a distintos niveles en el interior de la membrana y se hubieran deformado en el proceso de fractura.

Los hemidesmosomas, o semidesmosomas, constituyen un tercer tipo de unión adherente que se halla en las células epiteliales. Sirven como lugares de anclaje para los manojos de tonofilamentos, pero, a diferencia de los desmosomas puntiformes, no unen células adyacentes sino células epiteliales individuales con la matriz de tejido conectivo subyacente, impidiendo así que se separen las dos capas. Las células que se hallan sometidas a tensiones mecánicas fuertes, como las de la cerviz del útero, poseen un número extraordinariamente alto de desmosomas puntiformes y de hemidesmosomas. Estas uniones están interconectadas mediante redes de tonofilamentos que sirven para limitar la distensibilidad

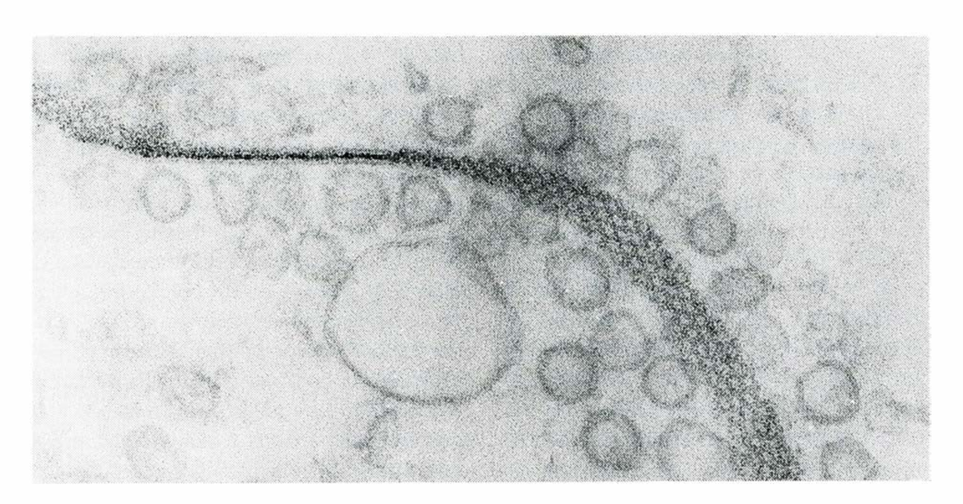
de las células y, por medio de los filamentos de unión transmembranal, para distribuir las fuerzas divisorias que actúan sobre células individuales en el conjunto del epitelio y en los tejidos subyacentes. De este modo, las uniones minimizan los efectos desgarradores de dichas fuerzas.

El desarrollo y el mantenimiento de la organización en el interior de un organismo depende del intercambio de información entre sus células constituyentes. Hay unas estructuras que son mediadoras de la transferencia directa de mensajes químicos entre células, que llamaremos uniones con hendidura ("gap junction"). En los últimos años, se ha investigado mucho en el campo de estas uniones, pues ofrecen un potencial interesante para la exploración de uno de los problemas fundamentales de la biología moderna: el control del crecimiento y diferenciación celular en los organismos superiores.

En cortes observados al microscopio electrónico, parece que el espacio intercelular se estrecha de su anchura normal, que son unos 250 angstroms, hasta unos 30 angstroms. A menudo, el espacio se encuentra obscurecido si los cortes no son lo suficientemente finos o si las membranas no se han teñido de manera adecuada; el resultado es que, durante varios años después del descubrimiento inicial de las uniones con hendidura, solían confundirse con uniones herméticas. Sin embargo, hoy ya está claro que los dos tipos de unión tienen características estructurales y funcionales muy diferentes.

Como demostraron inicialmente Jean-Paul Revel y Karnovsky, de la Facultad de Medicina de Harvard, las uniones con hendidura no obstaculizan el flujo de marcadores de metal pesado, como el hidróxido de lantanio, a través del espacio intercelular. Si se realiza un corte fino tangencial a la unión, se verá, sin embargo, que el microscopio electrónico revela una ordenación hexagonal de estructuras cilíndricas en las que el marcador no ha penetrado. En algunas ocasiones, el marcador llena un agujero central en estos cilindros, poniendo de manifiesto que se trata de la abertura de un pequeño canal. En las réplicas de congelación-fractura, las membranas plasmáticas adyacentes contienen, en la región de la unión con hendidura, una distribución en forma de disco de partículas poco distanciadas y se puede demostrar que las partículas de las dos distribuciones (pertenecientes a las dos semimembranas) están alineadas en el espacio intercelular.

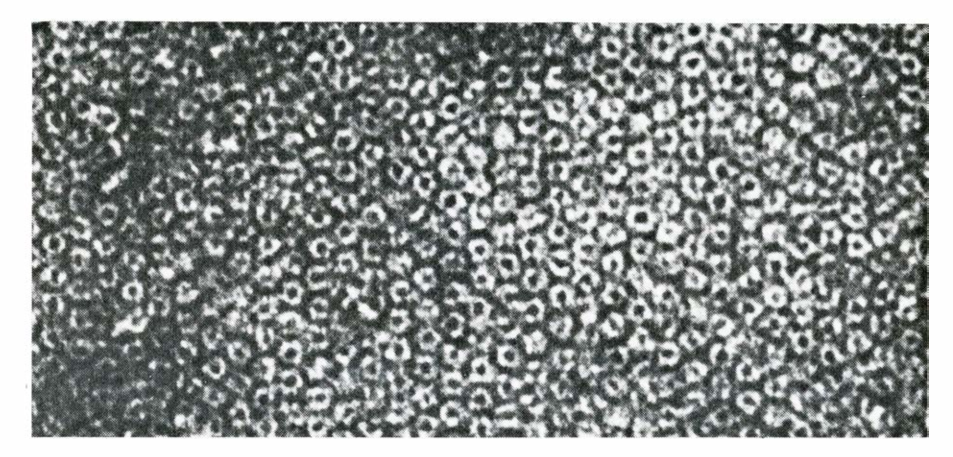
Las partículas de la unión con hendi-



CORTE ULTRAMICROTOMICO DE UNA UNION CON HENDIDURA, que revela aspectos de su organización estructural. Esta microelectrografía muestra una unión con hendidura de cangrejo de río, habiendo sido sumergido el corte en una solución de la sustancia opaca a los electrones hidróxido de lantanio. En la sección transversal vertical *(izquierda)*, el depósito de lantanio parece llenar el estrecho espacio intercelular. En la sección tangencial *(derecha)*, la sustancia opaca pone en evidencia una ordenación hexagonal de partículas ligeramente teñidas. La micrografía, proporcionada por Camillo Peracchia, de Rochester, tiene un aumento de 100.000 diámetros.



REPLICA DE UNA UNION CON HENDIDURA preparada por congelación-fractura; nos presenta una ordenación en forma de disco de partículas en la cara de la semimembrana citoplasmática de la célula inferior y una ordenación complementaria de hoyos en la cara de la semimembrana externa de la célula superior. Las partículas de la unión con hendidura terminan en el espacio intercelular, donde entran en contacto con el conjunto de partículas de la célula superior. (Este último conjunto de partículas fue arrancado durante el proceso de fractura.) Obsérvese el estrechamiento del espacio intercelular en la unión con hendidura. Aumento: x 100.000.



UNION CON HENDIDURA AISLADA, que ha sido teñida, y que aparece como una rejilla de partículas cilíndricas con el centro lleno de la sustancia de tinción. Estos centros actúan como canales intercelulares. La micrografía, realizada por N. B. Gilula, aumenta la unión 410.000 veces.

dura parecen formar conductos o canalillos intercelulares que harían de puente, salvando las membranas adyacentes y el espacio intercelular, permitiendo así el intercambio de moléculas entre células. Este modelo explica la posibilidad de que los marcadores de metal pesado se introduzcan en el espacio intercelular infiltrándose en la unión con hendidura. Si consideramos, por analogía, las dos células adyacentes como un par de depósitos metálicos y la unión con hendidura como un conjunto de pequeños conductos que las ponen en comunicación, resulta evidente que las sustancias externas a los depósitos serán capaces de pasar por los espacios entre los conductos, aunque no podrán penetrar en el interior de ellos ni tampoco dentro de los depósitos.

Los elementos básicos del modelo de la unión con hendidura fueron postulados a principios de los años sesenta por Werner R. Loewenstein, en la Facultad de Medicina y Cirugía de la Universidad de Columbia, basándose en unas mediciones eléctricas realizadas en células vivas. Loewenstein (como también, algo antes, Silvio Weidman, de la Universidad de Cambridge, Edwin J. Furshpan y David D. Potter, del University College London) había observado que, cuando se inyectaba una corriente eléctrica en el interior de una célula, provocando un cambio del potencial que existe de un lado a otro de la membrana, se podía detectar un cambio de voltaje de casi la misma magnitud en las células adyacentes. Loewenstein concluyó que las células debían poseer unos puntos de comunicación de baja resistencia eléctrica que permitieran que pasara el cambio de voltaje de unas células a sus vecinas, hipótesis que contradecía el concepto clásico de que las células eran entidades completamente autónomas. De hecho, se vio que si se inyectaban en el interior de una célula determinada moléculas de marcador de tamaño pequeño, como el colorante fluorescente amarillo de proción, el colorante aparecía en las células con las que la célula inyectada presentaba acoplamiento eléctrico.

Este hecho indujo a pensar que los canales de las uniones con hendidura que van de una célula a otra permitían el paso de determinadas sustancias entre células vecinas. La utilización de moléculas marcadoras fluorescentes de distintos tamaños permitió demostrar que las moléculas de peso molecular hasta de 1000 daltons podían pasar fácilmente de una célula a otra por los canales de conexión cuyo diámetro aproximado es de 20 angstroms. Dentro de este orden de tamaños caen los iones, la mayoría de los azúcares, los aminoácidos (elementos componentes de las proteínas), los nucleótidos (elementos componentes de los ácidos nucleicos), las vitaminas y las moléculas "mensajeras", como las hormonas esteroides y el AMP cíclico. Por tanto, las células conectadas entre sí por uniones con hendidura pueden extraer estas sustancias de un acervo intercelular de las mismas.

Es posible que el postulado intercambio de moléculas reguladoras en las uniones con hendidura tenga gran importancia en la regulación del crecimiento y de la diferenciación de grandes grupos de células durante el desarrollo embrionario. También parece que el acoplamiento sea vital para la distribución de los nutrientes en los embriones antes de que quede establecido el sistema circulatorio. Potter y Furshpan, como también Edwin S. Lennox, del Salk Institute for Biological Studies, demostraron que las células vitelinas del embrión de calamar están eléctricamente acopladas con todas las demás células del embrión hasta el momento en que se inicia la circulación sanguínea. Una vez ha llegado a ser funcional el sistema circulatorio, las células vitelinas tan sólo poseen acoplamiento con otras células vitelinas.

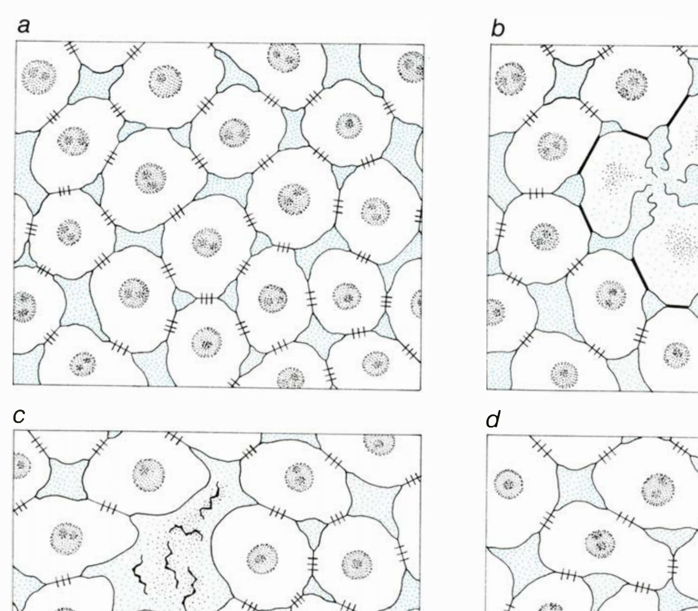
En los tejidos excitables por estímulos eléctricos, las uniones con hendidura sirven para transmitir las señales eléctricas y, debido a esta capacidad, se les denomina sinapsis electrotónicas. Estas uniones hacen posible que la actividad eléctrica de una célula se transmita a la célula adyacente sin que haya necesidad de la mediación de un neurotransmisor, o sustancia mensajera. Puesto que la transmisión eléctrica es prácticamente instantánea, las sinapsis electrotónicas se suelen encontrar en tejidos donde reviste considerable importancia la velocidad de la respuesta o la sincronización exacta de las actividades de un gran número de células. Por ejemplo, son uniones con hendidura las que sincronizan las contracciones de las células musculares del corazón y las contracciones de las células de músculo liso responsables de los movimientos peristálticos del intestino.

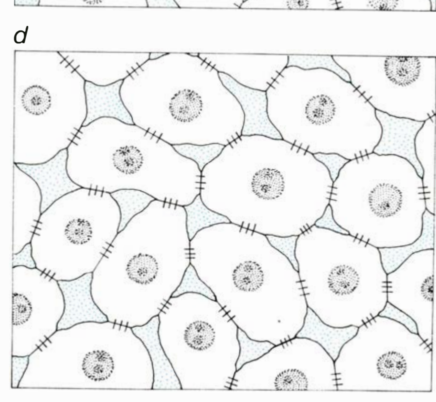
Otra función de las uniones con hendidura que resulta más exótica fue obser-

MODELO DE UNA UNION CON HENDIDURA, que representa los elementos estructurales que permiten el intercambio de nutrientes y de moléculas señalizadoras entre las células, sin que se pierda material en el espacio intercelular. Los "conductos" de comunicación están formados por partes de partículas colindantes, que a su vez constan de seis subunidades proteicas con forma de palanqueta, que ocupan toda la bicapa lipídica de cada membrana celular. El canal que pasa por medio de las partículas cilíndricas tiene un diámetro aproximado de 20 angstroms, limitando así el tamaño de las moléculas que pueden pasar a su través. A diferencia de la unión estrecha, los fluidos y los marcadores pueden penetrar en el espacio intercelular de la unión con hendidura.

vada por Michael V. L. Bennett, de la Facultad de Medicina y Cirugía, al estudiar el control nervioso del órgano eléctrico del pez sudamericano *Mormyrus,* el cual aturde a su presa con un choque eléctrico. En estos peces eléctricos, la fuerza del choque depende de la sincronización exacta de las descargas del gran número de células excitables que constituyen el órgano eléctrico, sincronización de la que son responsables unas motoneuronas de la médula espinal. Bennett encontró que, cuando estimulaba una de las células nerviosas motoras que inervaban el órgano, todas las demás neuronas inervadoras emitían una descarga simultáneamente, ya que el impulso se había transmitido de una neurona a otra mediante uniones con hendidura.

La permeabilidad de los canales de la unión con hendidura está estrictamente regulada por la concentración de iones calcio en el interior de la célula. Normalmente, suele ser bastante bajo el nivel de iones calcio libres en el citoplasma. Cuando se inyecta calcio en una célula, en poco tiempo esta célula pierde su acoplamiento con las vecinas. Si luego se deja que la célula siga su proceso propio, las mitocondrias intracelulares van retirando gradualmente el exceso de calcio hasta conseguir que los niveles vuelvan a ser normales y entonces los canales de unión reasumen su funcionalidad. Si inyectamos en una célula una sustancia sintética que se une a los iones calcio libres, como puede ser el ácido etilendiamintetracético (EDTA), también se restablece el acoplamiento eléctrico que había quedado bloqueado antes por la inyección de calcio.





REACCION INICIAL DE LAS CELULAS DAÑADAS: cierran sus uniones con hendidura en los límites de la herida, impidiendo así la pérdida de nutrientes vitales por parte de las células intactas. Esta reacción viene provocada por la entrada de iones calcio a la célula dañada, calcio que procede del líquido extracelular *(a, b).* Al cerrarse la herida, el acoplamiento se vuelve a establecer entre células intactas en el espacio de 30 minutos *(c, d).* Si las uniones con hendidura no se cerraran cuando se dañan las células, un tejido con muchas uniones con hendidura, como el estómago o el hígado, no sobreviviría a la destrucción de una sola de sus células. Esta ilustración está basada en otra elaborada por Werner R. Loewenstein, de la Universidad de Miami.

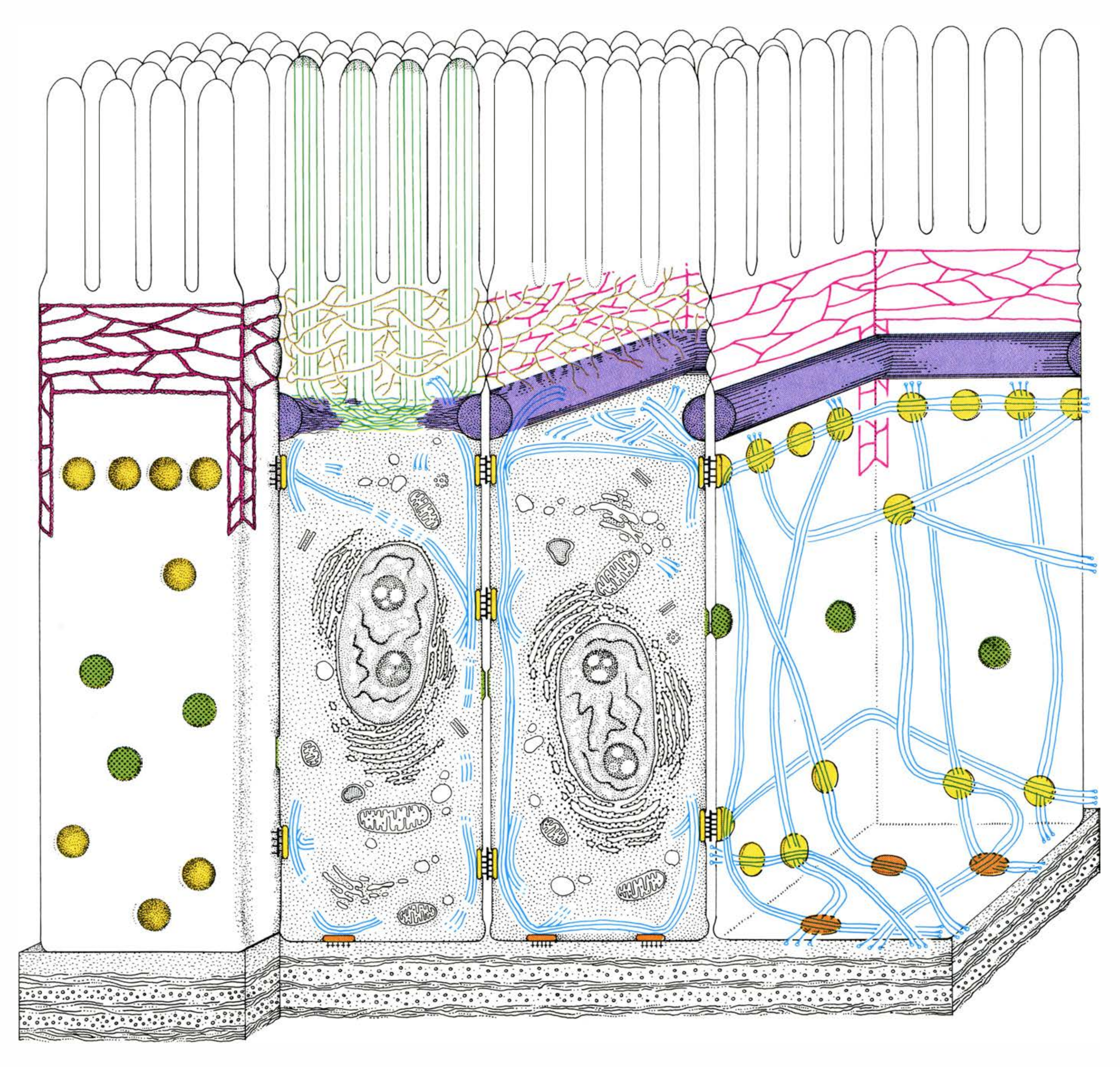
Loewenstein, que se halla actualmente en la Facultad de Medicina de la Universidad de Miami, ha sugerido la idea de que la sensibilidad de las uniones con hendidura a los iones calcio es esencial en relación con la capacidad de los tejidos para repararse tras una lesión o una herida. Si perforamos una célula, los niveles de calcio relativamente elevados en el medio circundante penetran en el interior de dicha célula, provocando el cierre de las uniones con hendidura formadas por la célula afectada. Como resultado de esta reacción, la pérdida de nutrientes por parte de las células vecinas no dañadas es mínima. Al irse cerrando la herida, las células que entran en contacto con otras células normales forman uniones con hendidura nuevas y funcionales en el espacio de 30 minutos. Camillo Peracchia, de la Facultad de Medicina y Odontología de la Universidad de Rochester, ha comprobado que, cuando se cierran los canales de las uniones con hendidura, disminuye el diámetro de las partículas que los forman, lo que sugiere que el cierre se lleva a cabo por "colapso" de los elementos de los canales.

Hoy en día se está registrando un especial interés en aislar y caracterizar químicamente las partículas componentes de las uniones con hendidura. Goodenough ha observado que las uniones con hendidura aisladas contienen principalmente una proteína, la conexina, cuyo peso molecular es de 18.000 daltons, y que está constituida por dos cadenas de unidades aminoácidas. Hace poco tiempo, en colaboración con D. L. D. Gaspar, Lee Makowski y W. C. Phillips, de la Universidad de Brandeis, ha examinado partículas aisladas de dichas uniones mediante microscopía electrónica y técnicas cristalográficas de rayos X; ha llegado a la conclusión de que están constituidas por seis subunidades en forma de palanqueta que se reúnen formando una estructura cilíndrica de unos 70 angstroms de diámetro con una simetría hexagonal y un canal central.

Todavía es muy poco lo que se sabe acerca de cómo se ensamblan, a nivel molecular, las uniones intercelulares. Uno de los enfoques que se le ha dado al problema ha sido aislar las uniones y analizar bioquímicamente sus partes componentes. Otros investigadores han buscado variaciones de las propiedades de las uniones intercelulares en células cancerosas, que crecen de manera incontrolada y se separan de sus vecinas dispersándose por todo el cuerpo. Aunque muchas de las células cancerosas parecen normales en cuanto a su acoplamiento intercelular, algunas tienen un bajo grado de ensamblaje, lo que hace pensar que una de las causas del crecimiento celular incontrolado podría ser la incapacidad para emitir o recibir moléculas señalizadoras a través de las uniones con hendidura. Este defecto puede surgir o bien del corto número de uniones con hendidura o bien de una alteración de la organización estructural de las partículas de estas uniones. Por otra parte, las células cancerosas que presentan uniones con hendidura normales, quizá padezcan de algún fallo en su capacidad para responder normalmente a las moléculas señalizadoras transferidas. De cualquier forma, parece estar claro que la salud de un organismo, al igual que la de una sociedad, depende de la buena

intercomunicación entre sus unidades constituyentes. Todo hace pensar que las uniones con hendidura constituyen la vía principal de comunicación intercelular.

En resumen, las funciones de las tres principales categorías de uniones intercelulares están estrechamente correlacionadas con su estructura. La unión estrecha o hermética está formada por dos membranas plasmáticas adosadas, unidas por filas entrecruzadas de proteínas totalmente membranales, que crean un cierre impermeable. Los desmosomas están compuestos por proteínas totalmente membranales y proteínas periféricas que tienden un puente sobre las membranas plasmáticas y ponen en comunicación los esqueletos internos de las células adyacentes. La unión con hendidura implica la participación de proteínas totalmente membranales organizadas formando unos tubos que constituyen canales de comunicación entre los citoplasmas de células adyacentes. La continuación de las investigaciones en torno a la arquitectura de las uniones intercelulares debería llevarnos a entender mejor el modo en que se produce la integración celular, dando lugar a formas vivientes con capacidades mucho más perfeccionadas que las de las células aisladas.



DISTRIBUCION DE LAS UNIONES en el epitelio del intestino delgado. La unión estrecha o hermética *(rojo)* forma una franja alrededor del borde superior de cada célula y está reforzada por una red de filamentos finos *(marrón)*. Inmediatamente por debajo se halla el desmosoma de banda *(morado)* que ciñe el interior de la membrana celular con manojos de filamentos contráctiles. A estos manojos se halla asociado un denso entretejido de filamentos denominados red de adherencia. La red deriva principalmente del esparcimiento de filamentos de los manojos de microfilamentos centrales *(verde obscuro)* que se prolongan hasta el interior de las microvellosidades y mantienen su forma. Por debajo de la red de adherencia, encontramos los desmosomas puntiformes *(amarillo)*, que parecen puntos de soldadura entre membranas de células contiguas. Luego vienen las uniones con hendidura *(verde claro)*, zonas a modo de parche de contacto intercelular íntimo. En la base de la célula, los hemidesmosomas *(naranja)* acoplan las células epiteliales con la matriz del tejido conectivo subyacente. Manojos de tonofilamentos *(azul)* se entrecruzan en el interior de la célula yendo de un desmosoma a otro, formando un armazón resistente a la tensión para el citoplasma. En resumen, las células adyacentes del epitelio están acopladas tanto mecánicamente (por los desmosomas puntiformes) como metabólicamente (por las uniones con hendidura) y separan los líquidos digestivos del intestino de los líquidos tisulares (por uniones herméticas).

Tecnología hidráulica romana

Las realizaciones de los romanos para recoger, mover y suministrar agua en cantidad no han sido igualadas en 1500 años. Los sistemas de acueductos, de un modo particular, presuponen muchas y notables proezas de ingeniería

Norman Smith

Los romanos han pasado a la historia con la fama de haber sido particularmente buenos en el campo de la ingeniería. Sus diversas obras hidráulicas demuestran que la fama está justificada. En efecto, sus realizaciones para la captación, traslado y suministro de agua (a menudo a grandes distancias y a pesar de numerosos obstáculos) no han sido igualadas durante los 1500 años, aproximadamente, que han transcurrido desde el ocaso del imperio romano.

Un buen lugar y momento para empezar a considerar las obras hidráulicas romanas es la Roma de finales del siglo primero. En el 97 d.C., un hombre nuevo se hizo cargo de la inspección de aguas de la ciudad. Era Sextus Iulius Frontinus, que había sido en otro tiempo (del 74 al 78 d.C.) gobernador de Britannia y era autor de libros de agrimensura y estrategia. Durante siete años, hasta el momento de su muerte, realizó una gran labor para poner un poco de orden y eficacia en un servicio de utilidad pública que había estado mal dirigido y abandonado durante años. En su obra *De aquis urbis Romae* puso por escrito sus experiencias en la administración del sistema de servicio público de aguas que fue el más dilatado del mundo antiguo.

A finales del siglo primero, Roma estaba abastecida por nueve acueductos. El más antiguo de ellos, el Aqua Appia, había sido construido hacía más de 400 años; los más recientes, el Aqua Claudia y el Aqua Anio Novus, se habían mantenido en servicio, o averiado, durante menos de 50.

La mayor parte del caudal y toda el agua de mejor calidad venía del valle del río Anio (del propio río y de manantiales). Las longitudes de los acueductos variaban desde 19 kilómetros hasta más de 80. En total, Frontino asumió la responsabilidad de unos 482 kilómetros de canales cubiertos *(specus)* con secciones que variaban desde unos 2,75 metros cuadrados hasta la elevada cifra de 12 metros cuadrados.

Es creencia popular que los acueductos romanos discurrían en toda su longitud por encima de arquerías. Esta idea es completamente errónea. Hasta donde era posible –y los romanos eran ingenieros eminentemente prácticos– los trazados de los acueductos, tanto en Roma como en cualquier otro sitio, seguían una pendiente regular a nivel del suelo o por debajo del mismo. El empleo de túneles, largas arquerías, altos puentes sobre valles fluviales, o tuberías para cruzar depresiones profundas, eran los últimos recursos de que se echaba mano ante condiciones adversas que no se podían afrontar de otro modo. Está perfectamente claro que los ingenieros romanos no escatimaban esfuerzos, literalmente hablando, para evitar cualquier compromiso con su norma básica. En el conjunto de acueductos que abastecían la ciudad de Roma, por ejemplo, sólo un 5 por ciento, aproximadamente, del kilometraje era llevado sobre puentes.

Los conductos superficiales y subterráneos eran más fáciles de construir y de mantener. El acceso para limpiar y reparar los conductos subterráneos se verificaba a lo largo de cada *specus,* a través de pozos o aberturas llamadas *putei.* La basura extraída de los acueductos se apilaba junto a los *putei;* los arqueólogos modernos han podido encontrar los conductos subterráneos, olvidados hace tiempo, localizando estos montones de piedra, grava, lodo y depósitos calcáreos.

Con objeto de mantener la pendiente requerida, cinco de los acueductos de Roma tuvieron que ser llevados sobre puentes unos cuantos kilómetros, las magníficas arquerías que aún hoy constituyen un rasgo tan dominante de la Campagna. La economía se impuso, y así los cinco conductos utilizaron sólo dos puentes. El Aqua Tepula y el Aqua Iulia fueron llevados por el puente construido en principio para sostener el Aqua Marcia; y el Aqua Anio Novus compartió el puente de su contemporáneo, el Aqua Claudia. En términos puramente constructivos estas ampliaciones no eran difíciles de levantar: se trataba de poner un nuevo canal (o dos canales) encima del ya existente. Para los canales del Aqua Tepula y del Aqua Iulia, sobre el Aqua Marcia, se empleó hormigón revestido de ladrillo, y ladrillo, revestido de hormigón impermeable, llevaba el Aqua Anio Novus por encima del puente del Aqua Claudia.

A largo plazo, los tramos elevados de los acueductos no lograron el éxito apetecido. Los testimonios escritos y arqueológicos revelan la necesidad de amplias y frecuentes reparaciones, que acarreaban largas interrupciones en el fluir del agua. Frontino comenta la avería derivada de "defectos de la construcción original". Por ejemplo, el Aqua Claudia (cuyas obras duraron 14 años) quedó terminado en el año 52, reparado en el 71, después de 10 años de uso y nueve de interrupción, reparado otra vez en el 80, y vuelto a arreglar en el 84. La prueba de estas obras de reparación se aprecia en los tramos del Aqua Claudia que aún perviven. Muchos de los arcos han sido reconstruidos de modo ramplón, con gruesas capas de ladrillo, teja y cemento que a menudo se prolongan varios palmos por los pilares que los sostienen.

Sigue en la obscuridad la causa fundamental del problema. Se concibe que la superposición de uno o de dos canales sobre un puente existente resultase excesiva para la resistencia de los arcos, determinando la apertura de las juntas. Cabe también la posibilidad de que, al aumentar la carga sobre los pilares, se produjesen con el tiempo diferencias de asiento y la consiguiente recarga de los arcos. Frontino, que se percataba de los efectos de la temperatura, observa que una ventaja de los canales subterráneos estriba en que "por no estar sometidos

al calor y a las heladas, son menos propensos a sufrir desperfectos".

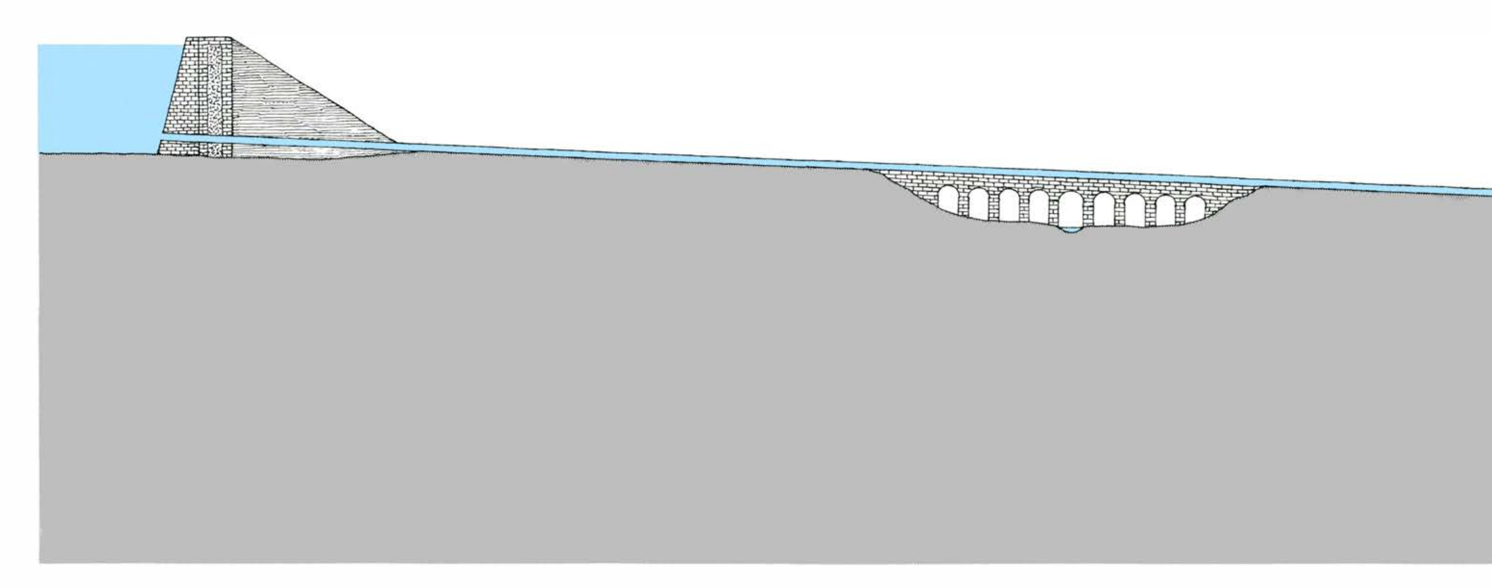
De todos modos, el germen de estos y de otros fallos estructurales eran las filtraciones, que junto con el robo de agua de los canales abiertos y de los tramos enterrados, si se podían alcanzar y perforar, ocasionaban importantes reducciones en el caudal de agua que, finalmente, se suministraba a los ciudadanos particulares y a las cisternas públicas de Roma. Esta era la cuestión que preocupaba a Frontino más que cualquier otra. Y al final lo derrotó, en buena parte porque él no era capaz de calcular las cantidades teóricas o prácticas del agua corriente. (Es notable que Frontino tuviese la impresión de que el volumen de la corriente sólo dependía del área de la sección transversal, y nada en absoluto de la velocidad. Si esta ignorancia estaba reducida a los funcionarios públicos es imposible comprobarlo ante la ausencia de una sola palabra que haya sobrevivido de un ingeniero hidráulico romano.)

¿Qué puede decirnos un cálculo moderno sobre la cantidad de agua suministrada a los habitantes de Roma, que en tiempos de Frontino ascendían quizás al millón? Hay que decir que nada; los datos disponibles no cubren las mínimas exigencias. Aunque se conozcan las áreas de las secciones transversales de los conductos, no se puede asegurar que alguna vez funcionasen llenos, pues su tamaño estaba determinado lo mismo por la necesidad de espacio para trabajar que por el espacio ocupado por el agua. La disposición y la capacidad de las tomas de los acueductos son desconocidas y lo mismo la cantidad de pérdida por filtración. Amén de todo eso, es seguro que sólo en raras ocasiones, si alguna vez lo hicieron, todos los acueductos de Roma estuvieron funcionando al mismo tiempo.

No obstante los defectos estructurales e hidráulicos que perturbaban continua-



ARQUERIAS de un puente portador de dos acueductos, a unos 10 kilómetros de distancia de Roma. La fotografía muestra parte del tramo más largo, de unos 1370 metros, comprendiendo 153 arcos. En este punto, el puente tiene una altura de unos 12 metros. La parte de ladrillo rota en lo alto llevaba el Aqua Anio Novus. Encima de los arcos discurría el Aqua Claudia por un canal de 0,93 metros de ancho y 1,86 de alto.



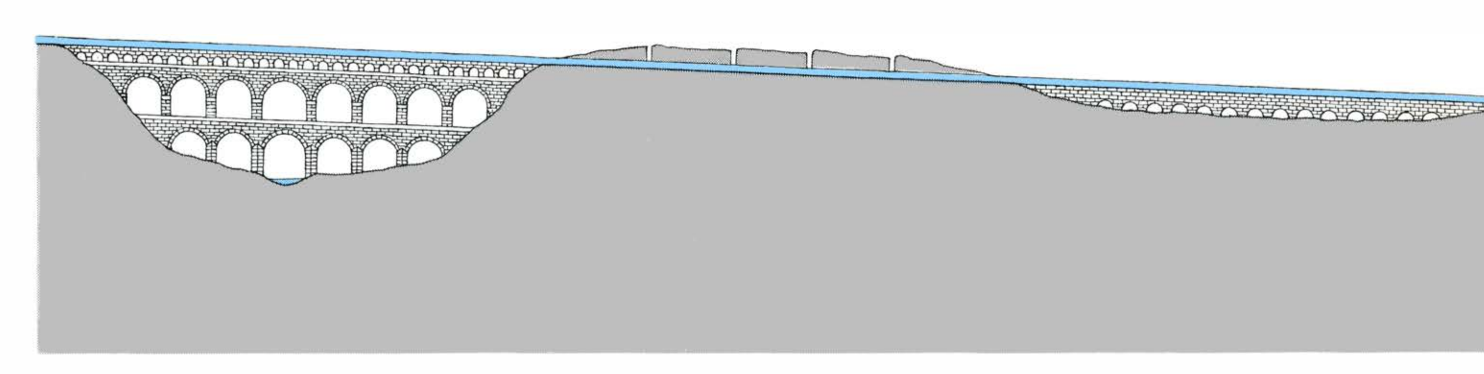
EL SISTEMA DE ACUEDUCTO típico, construido por los romanos, comprendía todos o la mayor parte de los elementos de ingeniería incluidos en este acueducto hipotético, que empieza en la parte superior izquierda de esta página y continúa hasta el fondo a la derecha de la página opuesta. En el manantial, el agua queda aprisionada detrás de una presa de tierra, revestida, en la cara correspondiente al agua, de

mente el funcionamiento del sistema y creaban dificultades para Frontino, la traída de aguas de Roma era una soberbia realización. El principio era una solución clásica a un problema fundamental de crecimiento de población urbana a gran escala. Frontino estaba tan impresionado, que llegado a un punto de su *De aquis* se siente impulsado a observar: "Con tal despliegue de edificios indispensables, portadores de tantas aguas compara, si quieres, las ociosas Pirámides, o las inútiles, aunque famosas, obras de los griegos". (El comentario es desconcertante porque las ciudades griegas habían sido las iniciadoras de la tecnología que Frontino consideraba tan beneficiosa para la vida urbana y tan sintomática del progreso civilizado.)

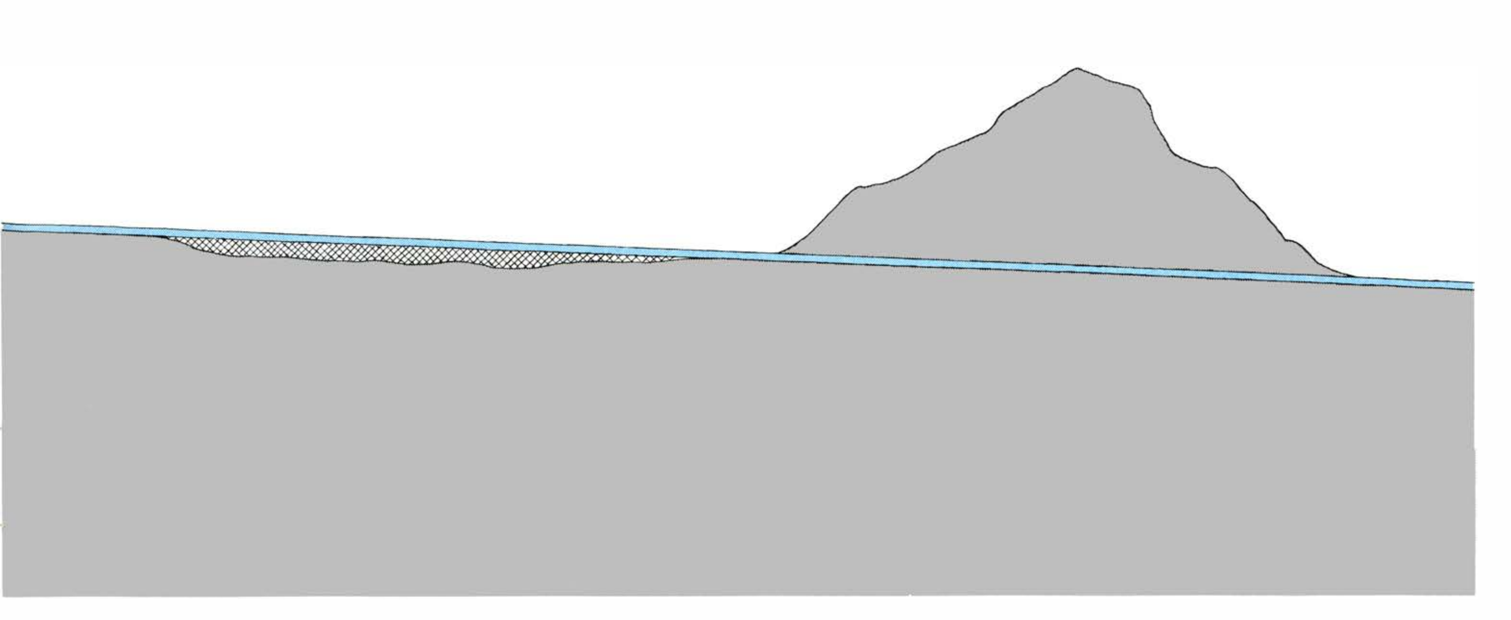
En cualquier evaluación del desarrollo de la ingeniería romana importa constatar cuánto fue asimilado de pueblos conquistados en general y hasta qué punto los ingenieros griegos en particular trabajaron para los romanos difundiendo así ideas y experiencias. No es menos cierto, sin embargo, que los ingenieros romanos adelantaron muchísimo más en la ingeniería del suministro de agua; a escala menor, el mismo servicio público era proporcionado en muchas ciudades de todo el imperio. Para este despliegue de capacidad en ingeniería civil fueron fundamentales dos importantes innovaciones romanas: la construcción de arcos y el desarrollo del hormigón hidráulico, que fraguaba rápidamente (incluso bajo el agua) y era fuerte e impermeable.

Se construyeron largas arquerías para llevar acueductos sobre terrenos llanos y hondonadas. Además de los existentes cerca de Roma, sobreviven hermosos ejemplos en el norte de Africa y en Mérida, en España. Por mucho que se estudiase el trazado desde el manantial a la ciudad, o por más que se considerasen las posibles opciones, los ingenieros romanos tenían que superar frecuentemente otros dos obstáculos topográficos: los montes y los valles de los ríos. Los terrenos altos eran superados generalmente por el procedimiento de circundarlos. La obra difícil, peligrosa y cara, de abrir túneles a largas distancias o grandes profundidades se eludía dentro de lo posible, como se haría en una empresa moderna. Se conocen pocos ejemplos de túneles de acarreo hidráulico romanos, entre otras razones por la dificultad de hacer coincidir dos segmentos de un túnel cuando éstos vienen en direcciones opuestas. (No debiéramos sonreír sin fundamento, porque lo mismo ocurre hoy en día, aunque raramente. El acierto en enlazar dos partes de un túnel largo todavía es, y con razón, una efemérides para celebrar con champán.)

En cuanto a los valles fluviales, la elección entre las dos soluciones disponibles venía impuesta por la profundidad del cruce. Hasta unos cuarenta metros se construía un puente. El magnífico Pont du Gard, cerca de Nîmes, en el sur de Francia, constituye un ejemplo sobresaliente. Sin embargo, y a pesar de su impresionante aspecto, el Pont du Gard es



LA CONTINUACION DEL ACUEDUCTO muestra cómo se transporta el agua, primero, a través de un valle bastante hondo (de 40 a 50 metros) por un puente de tres órdenes. A partir de allí, discurre algún tiempo bajo tierra por un túnel hecho por el procedimiento de cortar y tapar. Pozos llamados *putei* se abren a intervalos para la limpieza y reparación del túnel. Sobre un trecho de tierra llana el acue-



un muro de cantería y hormigón reforzado por una fila de contrafuertes de cantería. El agua es extraída por un túnel y discurre cierto trecho por un canal abierto. Cuando llega a un valle de poca profundidad, lo salva sobre un puente de arcos de cantería, cuya altura típica oscila entre los 25 y los 50 metros. En una depresión menos profunda continúa sobre un terraplén de tierra. Los romanos no solían construir túneles.

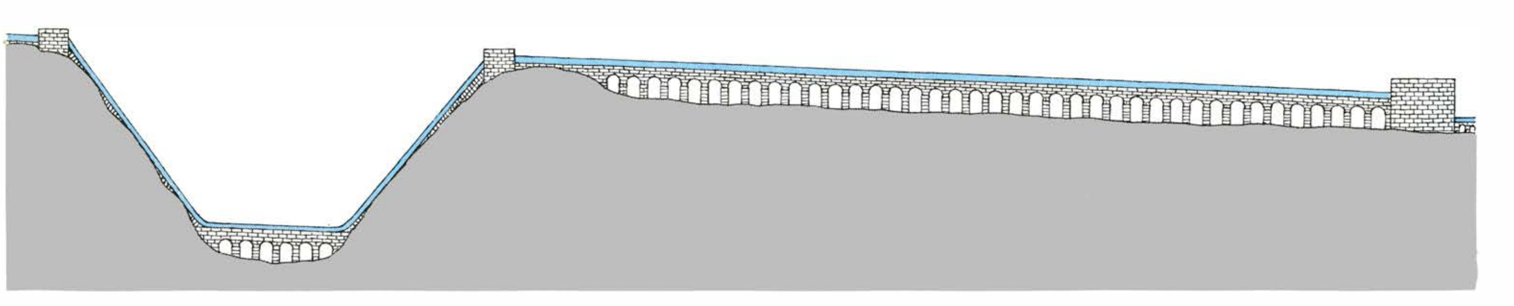
tosco de construcción. Los romanos alcanzaron su gran altura (unos 50 metros) por el primitivo recurso de superponer puentes, uno encima de otro, en tres niveles.

Gradualmente, los ingenieros romanos fueron perfeccionando sus métodos. En el siglo primero, se abandonó el sistema de superposición de puentes; posteriormente, el empleo del hormigón dio nuevo impulso al refinamiento y a la elegancia estructurales. En los acueductos de Mérida y de Cherchell, en Argelia, los altos pilares hechos de hormigón y revestidos de ladrillo y cantería son columnas continuas de arriba abajo, ligeramente acodaladas a intervalos, y sólo enlazadas firmemente en el extremo superior por los arcos que soportan el *specus*. Las arquerías de los acueductos romanos son un ejemplo muy instructivo del proceso de la evolución estructural.

Si un valle tenía una profundidad superior a los 50 metros, el puente por sí solo era inadecuado para sostener el acueducto. El coste de los materiales, cimbras, andamios, grúas, etcétera, resultaba prohibitivo. El solo peso del edificio terminado, sobre sus cimientos, podía ser también un factor disuasorio. La solución era emplear tuberías fabricadas laboriosamente con plomo o cerámica. Desde un depósito de cabecera situado al término del canal abierto por el que venía el agua, el sistema de tuberías (compuesto en general de varios tubos colocados uno al lado del otro) se hacía descender en un ángulo pronunciado por una ancha rampa de cantería. Una vez salvado el fondo del valle con un puente (que a menudo era de por sí un edificio muy grande), los tubos volvían a remontar la ladera opuesta sobre una rampa, hasta la altura de un depósito de recepción situado al comienzo de un nuevo tramo de conducción abierta. El depósito de recepción estaba situado en una posición más baja que el depósito de cabecera, lo bastante para garantizar el flujo adecuado del agua.

En Lyon, un conjunto de nueve tuberías de este género (su nombre técnico es el de sifones invertidos) servía a los cuatro acueductos que abastecían la ciudad. Los sifones principales variaban en longitud desde 800 metros hasta 4800, y alcanzaban profundidades de 60 a más de 120. Para soportar una presión tan alta, los sifones constaban de hasta diez tubos de plomo, de unos 25 centímetros de diámetro cada uno y de un espesor de pared de 18 milímetros. La cantidad de plomo (tal vez 15.000 toneladas) que tenía que ser extraído de la mina, fundido, transportado, fabricado e instalado era asombrosa. Sin embargo, se hacía, y el sistema funcionaba.

Hay un aspecto de los sifones que permanece obscuro, a saber: su diseño hidráulico y cómo se conseguía un grado de flujo satisfactorio. De algún modo, los ingenieros romanos dominaban la difícil cuestión de ajustar la capacidad de un sifón de varios tubos a la de los canales abiertos de conexión. Hay que sospechar que se llegaba de un modo



ducto cabalga sobre un terraplén bajo, perforado por arcos. Al llegar a un profundo cruce, un sifón invertido, consistente en tubos de plomo que parten de un depósito de cabecera, hace descender el agua por un lado, la pasa por un puente de arquería bajo, y la sube por el otro lado hasta un depósito de recepción. El acueducto sigue sostenido por una larga arquería hasta que llega al depósito de distribución en la ciudad.

empírico a la configuración requerida. Sabiendo por experiencia que un sifón tenía que descender un pie (30,5 centímetros) por cada 150 pies, aproximadamente, de su longitud, el ingeniero empezaría por construir primero el depósito de cabecera y las tuberías, y después recurriría a prudentes tanteos para situar al fin el depósito de recepción. En caso de necesidad, siempre se podía hacer un arreglo bastante tosco añadiendo o quitando un tubo. Existen pruebas de alguna operación de este género.

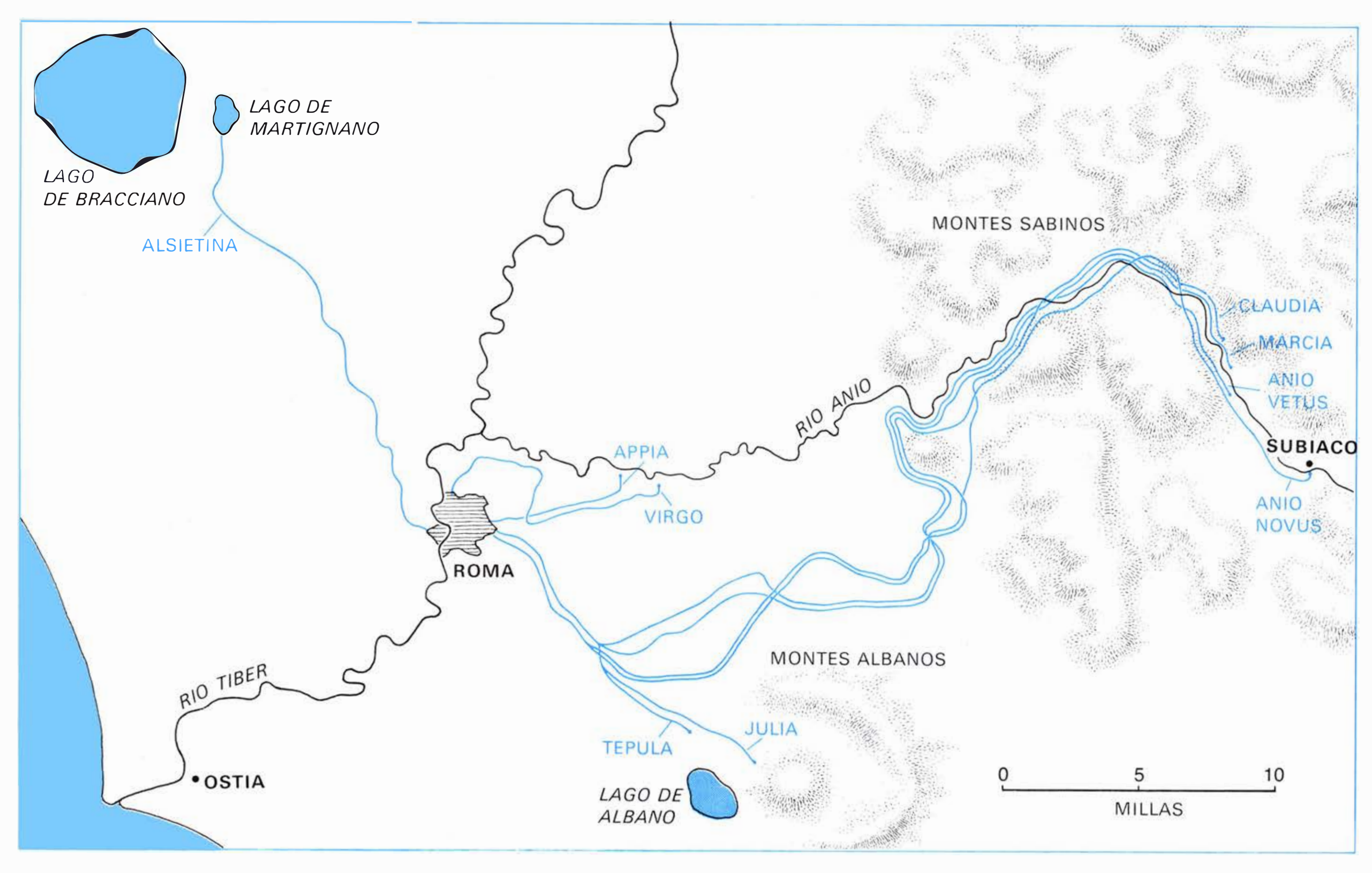
Tan importante como el copioso abastecimiento de agua potable a las ciudades era la cuestión de la calidad del agua. Como se puede imaginar, los sistemas aplicados eran de lo más rudimentario, considerando sólo factores tan evidentes como sabor, olor, aspecto y temperatura. Sorprende que la inmensa obra de construir el Anio Novus sólo consiguiese suministrar un agua turbia durante la época en que la toma se hacía del agua no limpia del río Anio. En el reinado de Trajano, y durante la gestión administrativa de Frontino, se logró una sustancial mejora trasladando la toma del Aqua Anio Novus a un depósito situado detrás de una inmensa presa en las cercanías de Subiaco. Es punto menos que imposible reconstruir con cierto detalle la presa en cuestión. Según las crónicas monásticas, la presa se derrumbó en 1305. Queda de ella una pintura medieval.

Algo se puede aprender de las realizaciones romanas en esta exigente rama de la ingeniería observando unas cuantas presas que aún se conservan. Es notable que dos de ellas presten servicio todavía. Las dos fueron construidas como depósitos de agua para Mérida. En esencia, se trata de diques de tierra de proporciones gigantescas, bien reforzados por dentro y por fuera mediante paredes y contrafuertes de cantería y hormigón. La supervivencia de estos notables edificios (que ahora prestan servicio con fines de regadío) se debe en gran parte a la conciencia, por parte de los ingenieros romanos, de la necesidad de aliviaderos, que son esenciales para que un dique de tierra no se vea barrido por la primera inundación que le rebose por encima.

El almacenamiento de agua detrás de grandes presas fue una innovación reveladora del compromiso de la sociedad romana con un abastecimiento fiable de agua, incluso en zonas de escasa pluviosidad. Las presas eran igualmente esenciales para el regadío. Una serie de presas romanas para el riego, en varios estados de reparación, se conservan en el Oriente Medio y Africa del Norte.

En términos agrícolas, el imperio romano comprendía tres regiones completamente distintas. Al norte de los Alpes, las tierras húmedas, pesadas y densamente arboladas de la Europa noroccidental planteaban problemas agrícolas que los romanos nunca resolvieron. En la cuenca del Mediterráneo, la mayor parte de los trabajos agrícolas podían llevarse a cabo sin irrigación. Más allá de esta cuenca se encontraban los antiguos centros de agricultura de regadío basada en los ríos. El más importante de ellos era el Nilo; sus aguas, y su singular sistema de irrigación de cuenca, hacían de Egipto "el granero de Roma".

Egipto era excepcional. En las demás regiones (al este y sur del Mediterráneo) las explotaciones agrarias del desierto requerían técnicas especiales que los romanos aprendieron de los nabateos, un pueblo muy capaz que floreció en el Negev hasta ser arrollado por los ejércitos de Trajano en el 106 d.C. Las principales características de la pluviosidad en el Neger, y a lo largo de la costa norte de Africa, son la intensidad y la breve duración de las precipitaciones. Las violentas crecidas que inundan las cuencas de los wadis unos cuantos días (o tal vez sólo unas cuantas horas) en el transcurso del año, llevan cantidades de agua de sorprendente magnitud y enormes concentraciones de fan-



ACUEDUCTOS DE ROMA, con su trazado a finales del siglo primero de la era cristiana. Llegó a haber nueve; el más antiguo de los cuales (Aqua Appia) había sido construido más de 400 años antes, y los más recientes (el Aqua Claudia y el Aqua Anio Novus) tenían alrededor de 50 años.

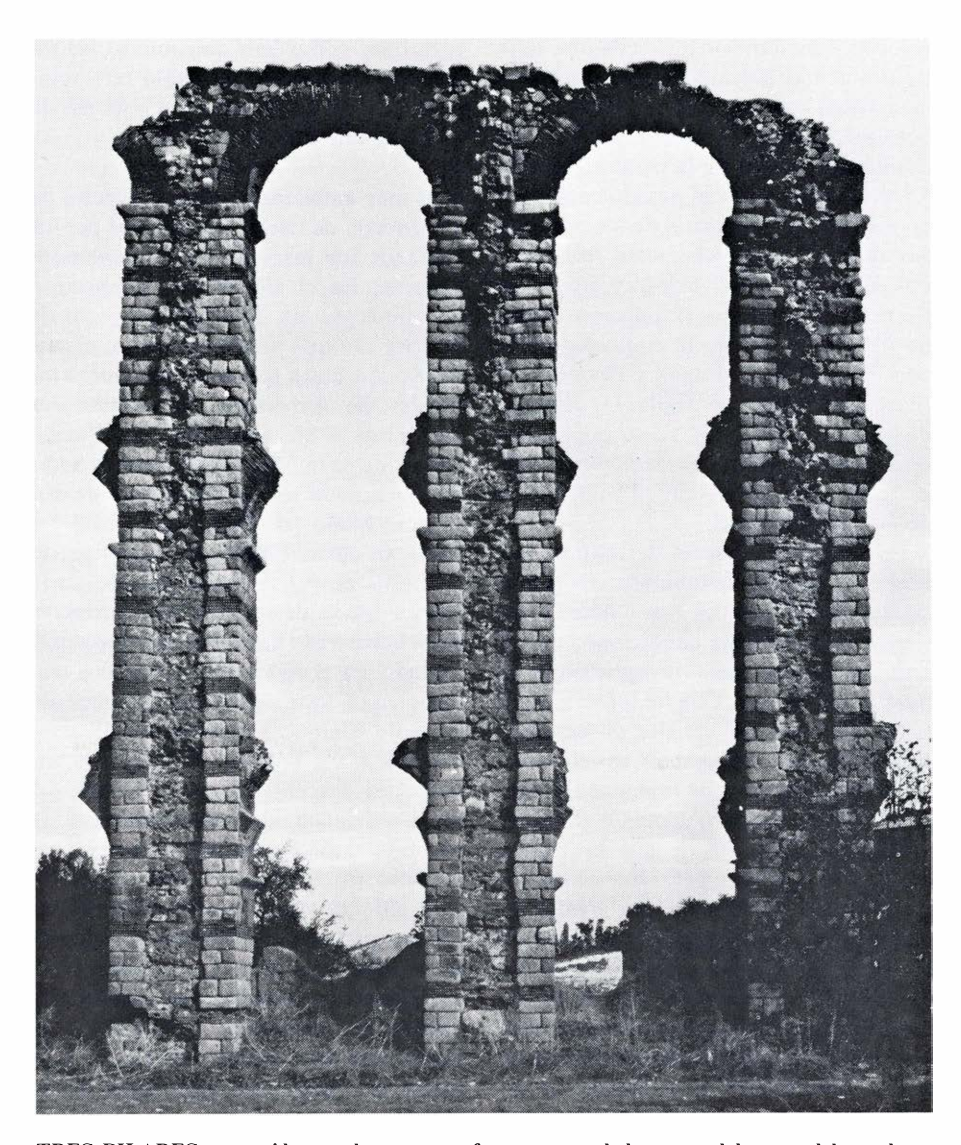
go. El arte de la irrigación de los wadis consiste en la retención eficiente de estas breves inundaciones para su empleo a largo plazo.

Uno de los procedimientos puestos en práctica utilizaba presas de desviación para dirigir el agua a tierras cultivadas o a depósitos de almacenaje. Otra posibilidad –y es en esta forma como la ingeniería se torna más compleja– consistía en terraplenar un wadi mediante una serie de presas cuidadosamente dispuestas, que con el tiempo iban quedando rellenas de fango a sus espaldas. Una vez que el wadi había quedado escalonado de esta forma, cada terraza podía sembrarse de cereales y cultivarse con regadío. Este método podía dar unos resultados agrícolas sorprendentes, aun cuando fuera empleado a la pequeña escala con que lo hacían los nabateos. Los romanos lo desarrollaron en grande, particularmente en el norte de Africa, con resultados impresionantes.

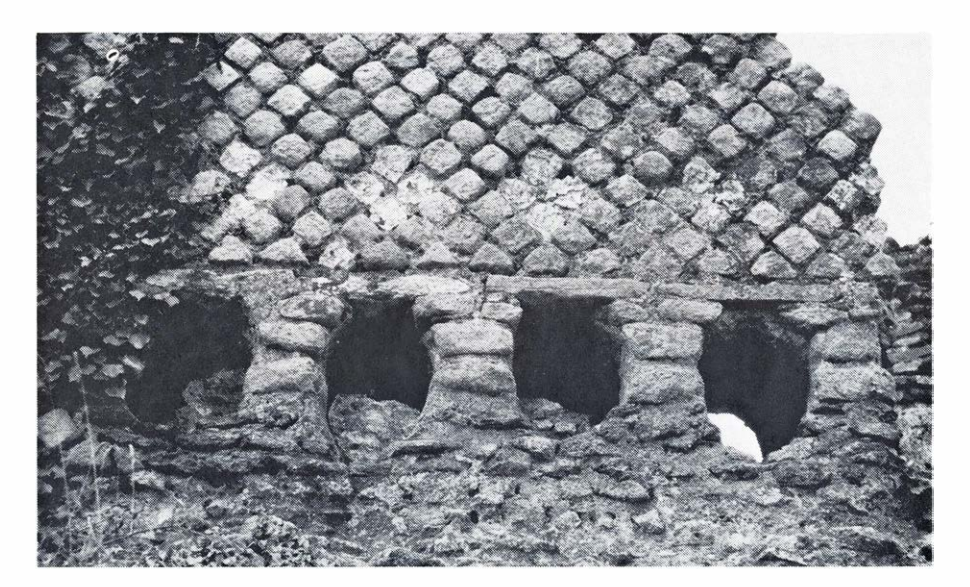
La construcción de los embalses de irrigación norteafricanos variaba según las condiciones dominantes. En los wadis abruptos y estrechos, con buenos cimientos de roca, la construcción típica constaba de un núcleo de cascajo y tierra, consolidado a veces con hormigón, que era revestido de bloques de cantería cuidadosamente cortados y ajustados; se sellaban las juntas con un fuerte mortero hecho de cal hidráulica. La cara de la presa que daba al agua era impermeabilizada con *opus signinum,* una especie de enlucido hecho con cal hidráulica mezclada con ladrillo o cerámica molida. La cara expuesta a la intemperie era escalonada y a menudo reforzada por contrafuertes para darle mayor estabilidad.

En wadis anchos, de pendiente más suave, las presas se levantaban con tierra en su mayor parte, es de suponer que forrada de cantería, para hacerlas impermeables y resistentes a la erosión. El abandono durante muchos siglos, y las inundaciones de los wadis, han removido prácticamente todos los restos de estas presas excepto los aliviaderos de cantería.

Las obras de irrigación de Africa del Norte ilustran el alcance de la construcción romana de presas. Los diversos materiales –cantería, tierra, mortero y hormigón– eran bien elegidos para su empleo; se utilizaban aliviaderos y compuertas para desviar el agua y almacenarla, generalmente para la agricultura, en ocasiones para el consumo doméstico. La obra ofrece un ejemplo impresionante de cómo un sistema de presas para regulación de inundaciones, retención de agua y conservación del suelo puede



TRES PILARES construidos por los romanos forman parte de los restos del puente del acueducto de Los Milagros en Mérida (España). En dos niveles intermedios se ven los restos salientes de dos arcos rebajados que fueron construidos para reforzar los pilares durante la edificación. La estabilidad longitudinal se logró cuando el canal y los arcos quedaron terminados encima del puente.



DEPOSITO DE CABECERA, residuo de un sifón del acueducto de Gier que abastecía Lyon, mostrando el tamaño de los tubos de plomo que llevaban el agua a través del sifón. El extremo de cada tubería, de las que este sifón tenía nueve, estaba encajado en uno de estos agujeros ovalados. En primer término se ve el comienzo de la rampa que llevaba los tubos por su pendiente hasta el valle. Este sifón media 1204 metros de longitud y alcanzaba una profundidad de 92 metros.

resolver simultáneamente toda una serie de problemas hidráulicos y agrícolas relacionados entre sí.

El más fundamental de los problemas de ingeniería civil de la península italiana, tan acuciante en el pasado como hoy en día, era la regulación de los ríos, con sus dos fines, estrechamente relacionados, de evitar inundaciones y drenar las tierras. En tanto que el ingeniero romano del Africa del Norte estaba logrando con éxito retener el agua y llevarla a la tierra, su colega de Italia se ocupaba de hacer lo contrario. Ya su predecesor etrusco había hecho ensayos para realizar esta tarea, la más difícil e intrincada de la ingeniería civil.

En muchas comarcas de Italia la pobreza del drenaje, combinada con la tendencia de los grandes ríos a inundar las tierras bajas, creaba condiciones que ni eran adecuadas para la agricultura ni para la habitación. Con frecuencia, el intento de construir canales de desagüe, algunos de ellos navegables, se veía coronado por el éxito. Los romanos trabajaron con firmeza para contener el imprevisible río Po, a fin de que las tierras arables no quedasen periódicamente empantanadas. El drenaje por medio de canales fue experimentado también en la costa del Tirreno, donde uno de los objetivos era la recuperación de las famosas Lagunas Pontinas. El problema era pavoroso; al final resultó insoluble, por lo menos para los romanos. La región era tan húmeda, baja y próxima al mar que, como han descubierto los ingenieros modernos, no había otra solución que la del bombeo, a una escala inimaginable para los romanos.

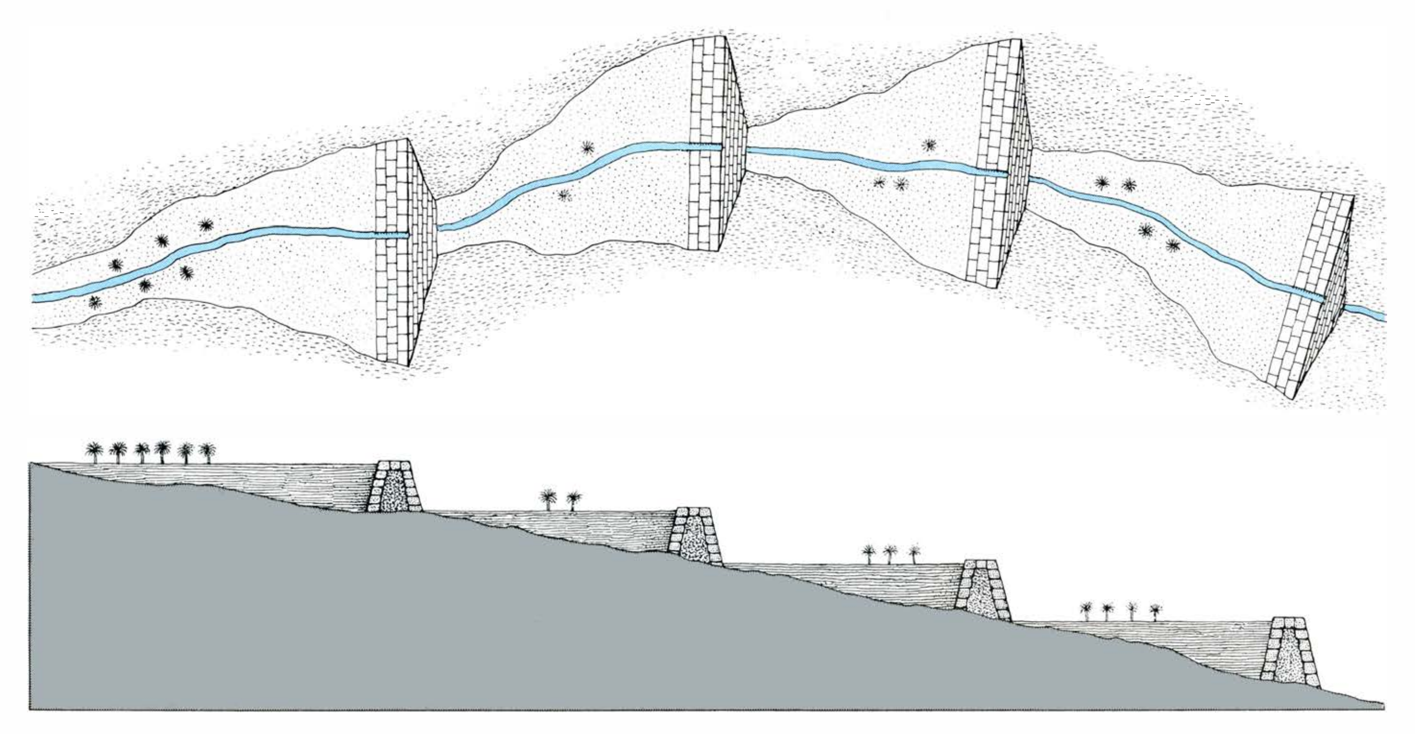
El más ambicioso de los proyectos de drenaje de tierras acometidos por los romanos fue iniciado por el emperador Claudio. En el año 41 d.C. empezaron las obras de un inmenso proyecto de perforación que posiblemente fue el más ambicioso hasta que se construyeron los túneles de ferrocarril. Su destino era dar salida a las aguas del lago Fucino (cerca de la moderna Avezzano), hacia un río situado a siete kilómetros de distancia. Plinio el Viejo, que visitó las obras en curso, quedó tan impresionado por este profundo túnel que se abría paso a través de la roca, que no recordaba haber visto nada más estremecedor. Cuando en el siglo XIX se volvió a trabajar en la obra, aparecieron tramos del tunel de Claudio y algunas modificaciones debidas a Trajano. Errores e inconsistencias descubiertas en el trazado del túnel indicaban que probablemente los romanos sacaron muy poco provecho del inmenso esfuerzo que habían realizado.

El enfoque de los problemas hidráulicos por parte de los romanos se centraba en la ingeniería civil. Los elementos de su ingeniería mecánica son mucho más difíciles de identificar, y escasas las reliquias que de la misma sobreviven. Son varias las cuestiones importantes suscitadas.

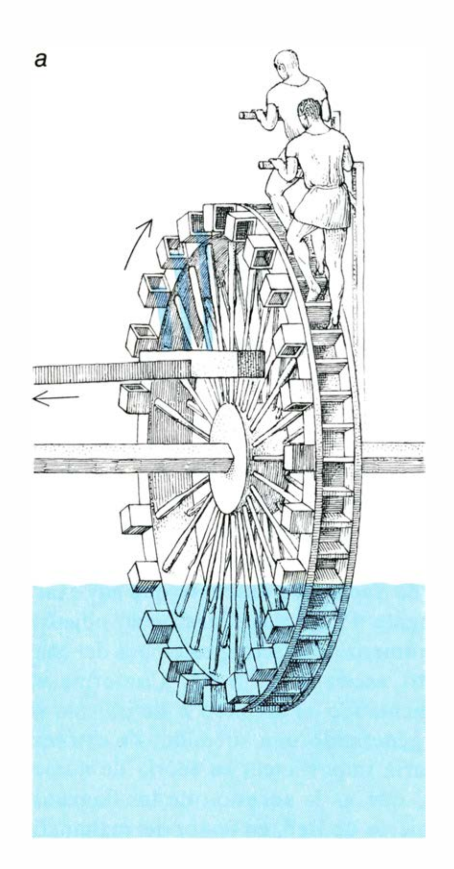
El origen, desarrollo y difusión final de todas las máquinas hidráulicas primitivas es sumamente obscuro, aunque parece natural suponer que la irrigación constituyó el estímulo inicial para la búsqueda de sistemas de elevación de agua. Una de las pocas fuentes de información de que disponemos es el tratado *De architectura* de Vitrubio, obra de gran amplitud sobre varias materias técnicas que se cree escrita alrededor del año 25 a.C.

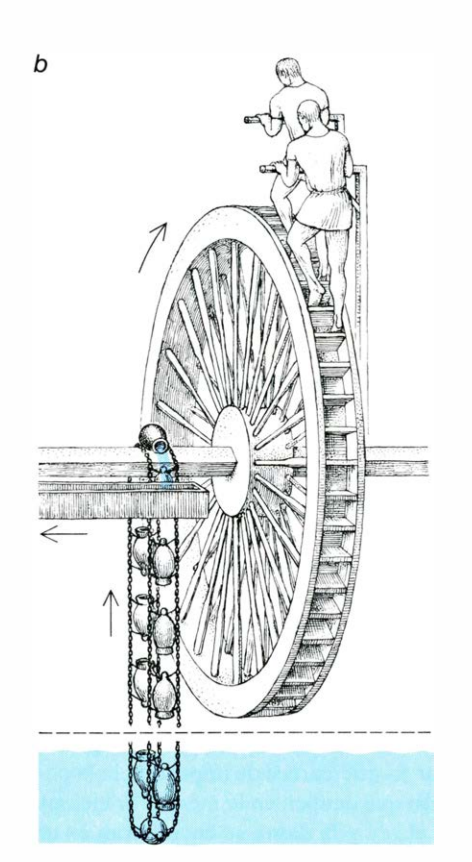
En este libro se encuentran descripciones objetivas de una serie de ruedas elevadoras de agua y propulsadas por energía hidráulica. Quiere decirse que su uso era corriente, pero no necesariamente extendido. Tres de los aparatos para elevar agua descritos por Vitrubio se relacionan íntimamente entre sí. La rueda de cangilones se instala en posición vertical y a su alrededor se ajustan recipientes; cuando la rueda gira, los cubos elevan agua a una altura aproximadamente igual al diámetro de la rueda. Para mayores elevaciones se recomienda una cadena de cangilones *(saquiyah).* En este caso, la rueda mueve una cadena sin fin de recipientes, que puede ser tan larga como lo permitan la fortaleza mecánica y la potencia del agente propulsor.

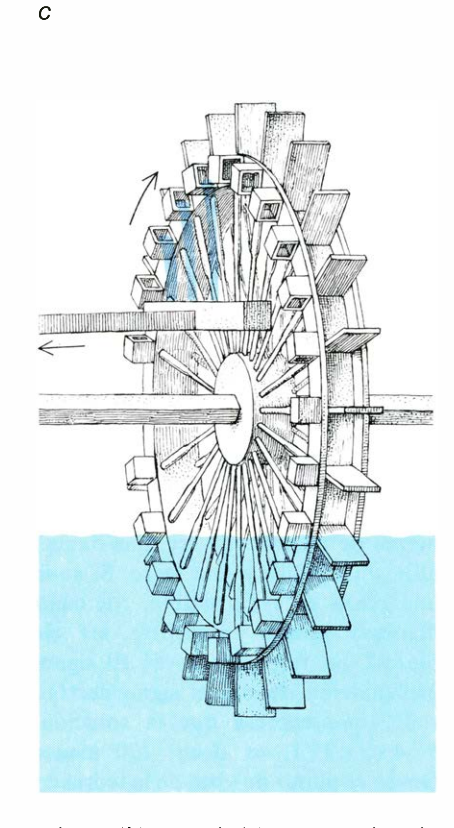
El tercer aparato es la *noria,* esto es, una rueda de cangilones, no movida por hombres o animales, sino por el empuje de la corriente de donde la rueda toma el agua. La descripción de Vitrubio es la primera referencia concreta a un sistema para obligar al agua corriente a realizar



IRRIGACION DE UN WADI en áreas desérticas. Un wadi es el lecho de un torrente que discurre poco tiempo, pero violentamente, durante las lluvias estacionales, y el objeto de la ingeniería es hacer posible el uso del agua durante largos períodos. Aquí se aprecia en plano *(arriba)* y en alzado *(abajo)* cómo los romanos escalonaban un wadi por medio de una serie de presas. Gradualmente el dorso de las construcciones se rellenaba de fango y el terreno formado así podía sembrarse y dar cosechas. El riego de éstas se efectuaba mediante inundación de las terrazas.







MAQUINAS ELEVADORAS DE AGUA usadas en las operaciones hidráulicas romanas. Comprendían la rueda de cubos o cangilones *(a)*, la *saquiyah,* o cadena de cangilones *(b)*, y la noria *(c)*, que era un ingenio accionado por agua. El origen de estos aparatos es todavía incierto.

un trabajo. Su pasaje sobre la *noria* va seguido de unas líneas sobre el molino de agua, una rueda, impulsada por la corriente, que acciona las ruedas del molino por mediación de un engranaje en ángulo recto. Los orígenes del molino de agua son obscuros. En particular, no sabemos si su punto de partida fueron las ruedas elevadoras de agua o viceversa.

Las máquinas giratorias no eran las únicas empleadas para elevar agua; los instrumentos de vaivén también entraban en juego, y es Vitrubio quien ofrece la descripción más antigua de una bomba de pistón. Dice que su relato se basa en la descripción, perdida hacía tiempo, del mecánico helenístico Ctesibius. En el 61 d.C. la misma máquina, de cilindros gemelos, accionada a mano, fue descrita por Herón de Alejandría como aparato extintor de incendios.

La ingeniería hidráulica de los romanos nos enfrenta con un vivo e intrigante contraste. Su repertorio de ingeniería civil era extenso y aplicado con largueza. Su entrega a la construcción, con todo lo que ésta implica en términos de materiales, mano de obra, tiempo y dinero, es impresionante a cualquier nivel. Por contra, su tecnología mecánica sólo fue objeto de un desarrollo marginal.

A fines del siglo I a.C., los romanos sabían cómo utilizar la energía hidráulica; y sin embargo, en lo que uno puede observar, prefirieron no hacerlo, o por lo menos no hacerlo con amplitud, y aun eso en una fecha avanzada. Para ojos modernos esta aparente distracción es enigmática y han sido muchos los intentos de explicarla. Probablemente la explicación ha de buscarse, en última instancia, en actitudes y expectaciones ante la tecnología que difieren radicalmente de las nuestras. En pocas palabras, la sociedad romana no podía simplemente comprender el concepto de una tecnología energética; no pensaba en términos de máquinas para ahorrar trabajo; y a causa de la naturaleza de la economía antigua, no sentía la necesidad de la producción mecanizada. Conviene recordar que para comprender con exactitud, y valorar objetivamente, cualquier parte de la tecnología antigua, hay que tratarla en este contexto, dispuestos a aceptar las actitudes y objetivos de la época, y resistiendo el impulso de imponerle los modernos.

El sistema romano de abastecimiento de aguas se revela como una de las grandes realizaciones de la Antigüedad. Con el declinar del imperio, tanto la autoridad administrativa necesaria para desarrollar y mantener tales sistemas, como el concepto mismo de obras públicas, se perdieron. Había de pasar mucho tiempo antes de que el concepto romano pudiera restablecerse una vez más.

Aún cuando en los siglos XVI y XVII algunas ciudades europeas trataron de conseguir más agua, lo hicieron con un enfoque distinto. Toledo, que había estado abastecida por un hermoso acueducto romano, terminó fracasando en sus esfuerzos por bombear el agua del Tajo. París consiguió bombear el Sena, y Londres el Támesis; pero ninguno de los dos ríos podía proporcionar agua suficiente por medios mecánicos. Además, el agua era muy inferior en calidad cuando no era verdaderamente mortífera. La ingeniería mecánica no era la solución. Sólo el New River de Londres, un canal abierto construido en 1613 para traer agua de una corriente situada a 65 kilómetros, representaba una solución al estilo romano.

Hace un siglo, poco más o menos, las poblaciones urbanas estaban creciendo a un ritmo sin precedentes, especialmente en Inglaterra. Brotes de cólera, en particular en Londres, llamaron la atención rápida y agudamente hacia el más crítico de todos los problemas urbanos de la era victoriana: la necesidad de abastecer de agua en una escala y de una pureza rara vez consideradas y nunca llevadas a efecto durante 1500 años. La solución fue resucitar el aparato entero de la tecnología de abastecimiento de aguas creado por los romanos.

Juegos matemáticos

Los versátiles números de Bell permiten contar desde particiones de un conjunto dado hasta versos monorrimos

Martin Gardner

Supongamos que en una mesa, dispuestos en fila, haya cinco platillos, marcados *a, b, c, d* y *e*. Supongamos también que se tengan cinco piezas de ajedrez: un rey, una dama, un alfil, un caballo y una torre. Si se sitúa una pieza en cada platillo, ¿de cuántas maneras pueden repartirse las cinco piezas? La respuesta es 5! El signo de exclamación denota el signo de "factorial", que expresa que la solución es $5 \times 4 \times 3 \times 2 \times 1$, es decir, 120 maneras. Desde el punto de vista de la teoría combinatoria, el problema es equivalente al de contar de cuántas formas pueden permutarse entre sí las letras *a, b, c, d,* y *e*. El resultado general, cuando hay n objetos, es $n!$

Modifiquemos un poco las reglas, permitiendo ahora situar un número cualquiera de piezas, de cero a cinco, en cualquiera de los platillos. ¿De cuántas formas diferentes podrán colocarse las piezas en los platillos? Como es evidente, una sola pieza puede colocarse exclusivamente en un único platillo de una sola forma. Dos piezas pueden distribuirse de cuatro maneras en dos platillos, según se muestra en la parte izquierda de la ilustración de la página opuesta. Experimentando con tres platillos y tres piezas se descubre que hay 27 disposiciones posibles. Dado que 1, 4 y 27 son respectivamente iguales a 1^1, 2^2 y 3^3 podría conjeturarse que hay n^n formas de distribuir n objetos en n platos. Conjetura que resulta efectivamente correcta. Las cinco piezas pueden colocarse en los cinco platillos de 5^5, o sea, 3125 maneras diferentes.

Haciendo la distribución de acuerdo con las nuevas reglas al repartir n objetos en k platillos resultan k^n disposiciones de los objetos. Por ejemplo, dos objetos pueden colocarse de 3^2, o sea, de 9 formas distintas en tres bandejas, como se muestra en la parte derecha de la ilustración de la página adyacente. No es difícil comprender por qué la fórmula anterior da los valores correctos. Hay k maneras de colocar el primer objeto en los k platillos. También el segundo objeto puede ir a cualquiera de los k platillos, por lo que también puede ser colocados de k maneras. Dado que hay n objetos, es evidente que pueden ser colocados en k platillos de $k \times k \times k \times \ldots \times k = k^n$ maneras.

Examinemos ahora un problema más difícil. Seguimos teniendo las mismas cinco piezas y los cinco platillos; la diferencia está en que ahora los platillos se consideran idénticos e indistinguibles, por lo que carece de importancia la posición que ocupen en la mesa. Por ejemplo, si el rey y la dama se encuentran en una bandeja y las otras tres piezas restantes en otra, no tiene importancia cuáles sean los platillos que reciben cada grupo de piezas. Todas las particiones del conjunto que sitúen al rey y la dama en el mismo platillo, y al alfil, caballo y torre en otro, sin que importe cuáles sean los utilizados, se considerarán idénticas, y contarán como una única forma de distribuir las plazas. ¿Cómo contar ahora todas las maneras de colocar los cinco objetos?

Como ocurría antes, es evidente que un objeto solamente puede ir a un platillo de un modo único. Dos objetos pueden situarse en dos platillos de dos maneras: o bien ambos en el mismo platillo o bien poniendo un objeto en cada uno de ellos. Muchas situaciones reales se adaptan a este modelo. Por ejemplo, marido y mujer tienen dos formas esencialmente distintas de ocupar camas gemelas idénticas: durmiendo en camas distintas, o juntos en la misma cama. Hay dos formas distintas en las que un policía puede esposar a dos detenidos: esposándolos por separado o encadenando uno con otro. La ilustración superior de la página 98 muestra los cinco modos posibles de situar tres objetos en tres platillos indistinguibles entre sí. A este modelo se ajustan las cinco formas en que tres personas pueden ocupar tres camas, en que tres naciones pueden formar alianzas, etcétera.

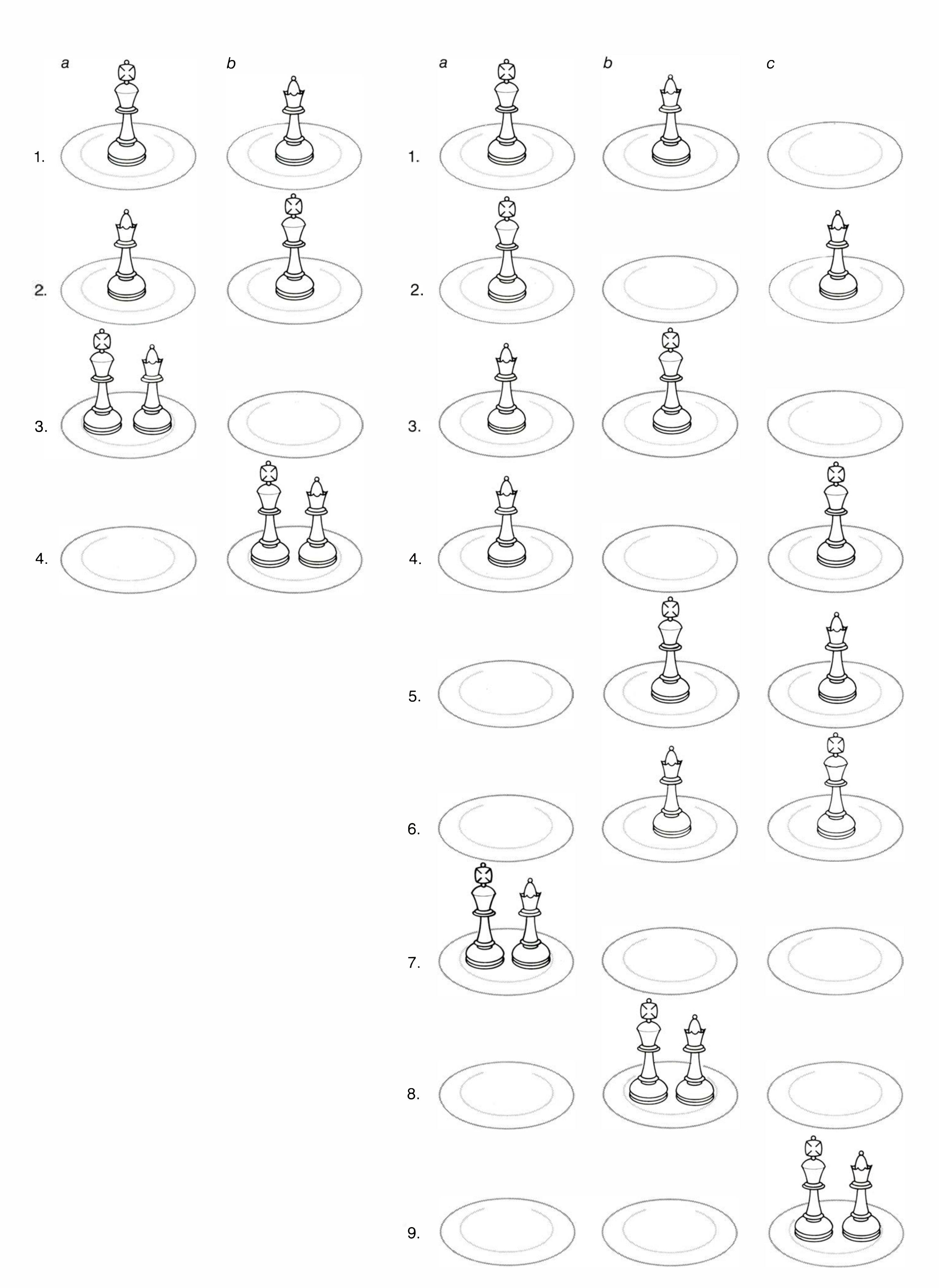
A modo de experimento, podríamos entretenernos ahora en contar efectivamente todas las formas de colocar cuatro objetos en cuatro platillos idénticos. Con terminología más técnica, el problema consiste en hallar de cuántas maneras puede un conjunto de cuatro elementos distintos descomponerse en subconjuntos no vacíos. Se descubre que hay exactamente 15 formas. Para cinco objetos, el número de descomposiciones del conjunto asciende ya a 52. Conforme va aumentando el número n de objetos se va generando una sucesión, de extraordinaria importancia en teoría de números, que es la sucesión de los llamados números de Bell, en honor del matemático americano –aunque de origen escocés– Eric Temple Bell, quien falleció en 1960. Son números íntimamente relacionados con los números de Catalán.

A pesar de que los números de Bell fueron conocidos desde mucho antes de que Bell se ocupara de ellos, fue él quien por primera vez los analizó profundamente y puso de manifiesto su importancia. En su primer trabajo sobre tales números, Bell explicaba cómo se despertó su interés por ellos. Había observado un error en un manual que daba el llamado desarrollo de MacLaurin de la expresión e^{e^x}, siendo e el trascendente (y trascendental) número de Euler, y x un entero positivo cualquiera. El desarrollo correcto es:

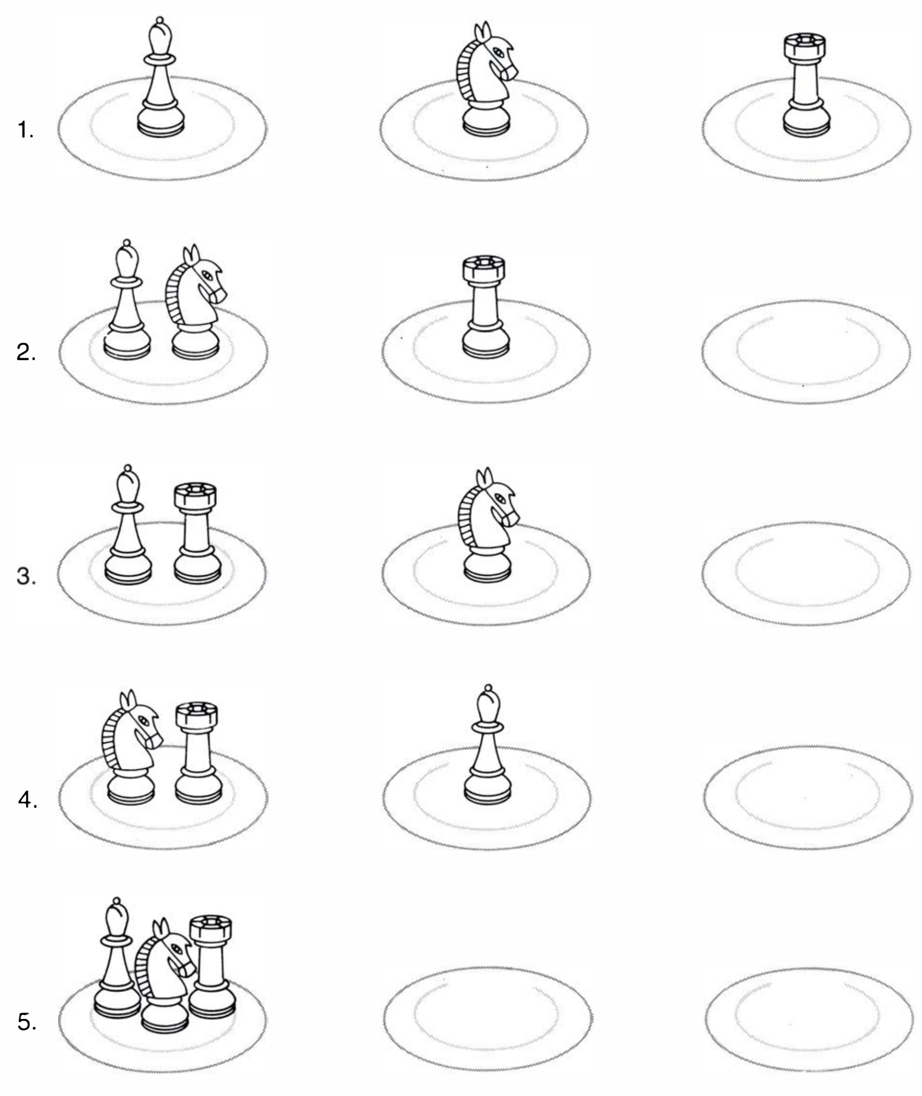
$$e\left(1 + \frac{x}{1!} + \frac{2x^2}{2!} + \frac{5x^3}{3!} + \frac{15x^4}{4!} \cdots\right)$$

Nótese que los coeficientes de las potencias de x son, precisamente, los números de Bell. (Bell denominó números exponenciales a tales números, pero después de que John Riordan, especialista en combinatoria, comenzara a denotarlos B con el fin de honrar a Bell, tales números pasaron rápidamente a ser conocidos por números de Bell.) Partiendo del desarrollo de MacLaurin resulta factible deducir la llamada fórmula de Dobinski para el n-ésimo número de Bell, B_n.

$$B_n = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}$$



Formas de colocar dos objetos en dos platillos distinguibles (**izquierda**) *y en tres platillos distinguibles entre sí* (**derecha**)



El tercer número de Bell ($B_3 = 5$) indica de cuántas formas se pueden colocar tres objetos en tres platillos indistinguibles

B_0	1
B_1	1
B_2	2
B_3	5
B_4	15
B_5	52
B_6	203
B_7	877
B_8	4,140
B_9	21,147
B_{10}	115,975
B_{11}	678,570
B_{12}	4,213,597

1						
1	2					
2	3	5				
5	7	10	15			
15	20	27	37	52		
52	67	87	114	151	203	
203	255	322	409	523	674	877
877	. . .					
⋮						

1	2	5	15	52	203	877	. . .
1	3	10	37	151	674	. . .	
2	7	27	114	523	. . .		
5	20	87	409	. . .			
15	67	322	. . .				
52	255	. . .					
203	. . .						
⋱							

Los números de Bell (**izquierda**) *y dos formas del triángulo de Bell* (**derecha**)

Fundamentalmente, Bell era un especialista de teoría de números, pero su fama se debe sobre todo a su clásica exposición de la historia de la matemática *The Development of Mathematics,* a su *Men of Mathematics,* más dirigida hacia el público en general, y a otras obras. Quizá los matemáticos de la joven generación queden sorprendidos al enterarse de que durante los años 20 y 30 Bell fue prolífico escritor de ciencia-ficción, con el pseudónimo John Taine. Cinco de sus novelas han sido reeditadas por Dover paperbacks, en dos volúmenes titulados *Seeds of Life and White Lily* y *The Time Stream.* En 1951 un número dominical de un diario de Pasadena presentó una reseña crítica de un libro de Bell, *Mathematics, Queen and Servant of Science,* realizada por John Taine. "La solapa posterior de la sobrecubierta", escribía Taine, "afirma que Bell es, seguramente, el máximo intérprete de las matemáticas. El recensor, que conoce bien a Bell, está de acuerdo."

Volvamos a los números de Bell, que podrían llamarse Bells e incluso Temple Bells. Los primeros trece números de Bell se muestran en la parte izquierda de la ilustración inferior de esta página. Por convenio, B_0 es igual a 1. Como puede verse, los números crecen con ritmo exponencial, o como expresó Poe en "The Bells", suben y suben "más y más y más alto, con ansia desesperada". El centésimo número de Bell consta de 126 dígitos.

Las fórmulas que expresan explícitamente el n-ésimo número de Bell son complicadas y difíciles de utilizar directamente para el cálculo de los términos de la sucesión, pero por fortuna hay un sencillo procedimiento recurrente que los proporciona fácilmente. El método se comprende mejor examinando el proceso de formación del triángulo de números de la parte superior derecha de la ilustración inferior de esta página. (Atendiendo una sugerencia del corresponsal Jeffrey Shallit, este triángulo se denominará triángulo de Bell.) Se comienza escribiendo un 1 en la parte superior, y otro 1 bajo el anterior. Como 1 y 1 suman 2, se escribe un 2 al final de la segunda fila. Se abre ahora la tercera fila con un 2. La suma de 2 y el número situado encima de él es 3, por lo que se sitúa el 3 a la derecha del 2. La suma de 3 con el guarismo situado sobre él es 5. Se coloca un 5 a la derecha del 3. Se prosigue de forma análoga, observando las dos reglas siguientes: El último número de cada fila es igual al primero de la fila siguiente; todos los demás números se

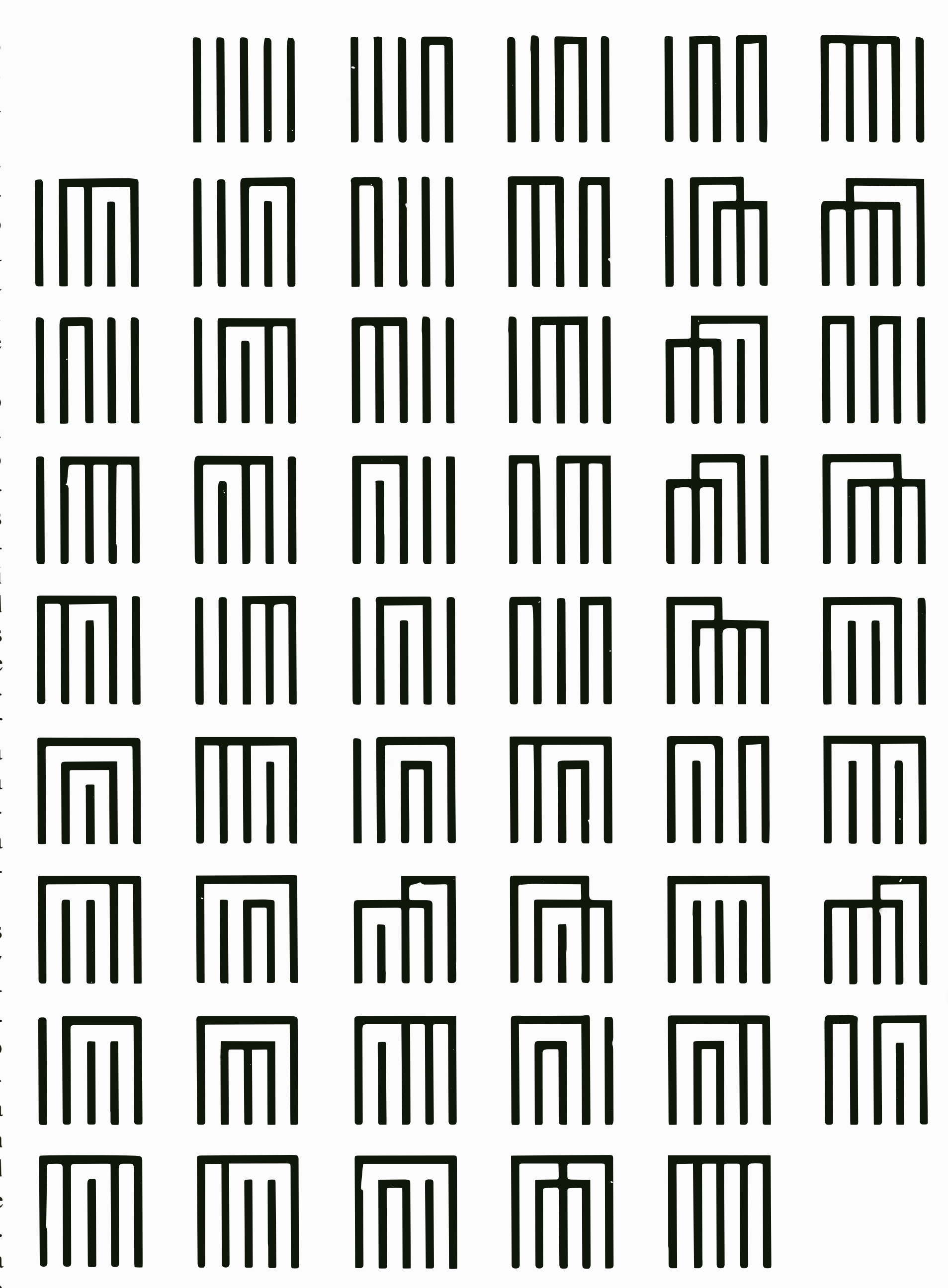
obtienen sumando el vecino izquierdo del número deseado con el número situado directamente sobre dicho vecino. La sucesión de números de Bell aparece en dos lados del triángulo. Girando convenientemente el triángulo, éste se convierte en un triángulo de diferencias, como se muestra en la parte inferior derecha de la ilustración inferior de la página opuesta; en él cada número, a excepción de los de la primera fila, es diferencia de los dos situados sobre él.

Lo mismo que el conocido triángulo de Pascal, el triángulo de Bell abunda en interesantes propiedades. En el triángulo de Bell que se muestra en la esquina derecha de la ilustración, las sumas de las filas horizontales resultan ser los números de la segunda diagonal infinita. Si se le suma al número de Bell situado al extremo de una fila, la suma de todos los números de ésta resulta el siguiente número de Bell. Sustituyendo cada número par por una *P* y cada número impar por una I, es fácil observar que cada tercer número de Bell es par. Así que la razón de números de Bell impares a números de Bell pares es de 2:1, y la suma de cada tres números de Bell consecutivos ha de ser par.

Se han observado docenas de curiosas propiedades de la sucesión de Bell, y los descubrimientos continúan. Por ejemplo, la "congruencia de Touchard" enuncia: $B_{p+k} \equiv B_k + b_{k+1}$ (módulo p), siendo p un número primo. Dicho de otra forma, si el subíndice n de B_n se expresa como suma de un número primo p y un número k, y si se divide B_n entre p, el resto obtenido será igual al resultante de dividir entre p la suma B_k y B_{k+1}. Tomando casos con k igual a cero, la congruencia enuncia que $B_p \equiv 2$ (módulo p). Dicho de otra forma, todos los números B_n en los que n sea primo dan resto 2 al dividirlos por n. Por ejemplo, el número B_{13} es igual a 27.644.437; al dividir este número entre 13 el resto es igual a 2.

Los números de Bell representan un importante papel en la teoría de números primos porque permiten contar cuántas descomposiciones factoriales admite un número, conocidos sus factores primos.

Por ejemplo, 30 tiene tres factores primos distintos: 2, 3 y 5. B_3 es igual a 5. Las cinco descomposiciones factoriales de 30 son $2 \times 3 \times 5$, 5×6, 3×10, 2×15 y 30. Es fácil comprender que este problema es isomórfico al de colocar tres objetos distintos en tres platos no distinguibles. Obsérvese que tres de los primeros 10 números de Bell son primos: 2, 5, 877. ¿Existen otros números de Bell que sean primos? ¿Existe un número de Bell primo máximo? No conozco la respuesta de ninguna de estas cuestiones.



Los 52 diagramas de Lady Murasaki, tomados de **The Tale of Genji**

Una de las sorprendentes aplicaciones de los números de Bell es que permiten contar el número de posibles esquemas de rimado de una estrofa poética. Una estrofa de un solo verso tiene solamente un posible esquema de rimado; una estrofa de dos versos tiene dos esquemas (los versos riman o no riman); una estrofa de tres versos tiene cinco esquemas (*aaa, aab, aba, abb* y *abc*), y así sucesivamente. Se dice que este "rhyming and the chiming of the bells" (Poe) fue observado por primera vez por el británico J.J. Sylvester, pero no he sido capaz de encontrar ninguna referencia al respecto en su librito *The Laws of Verse*. Quedaré muy reconocido a los lectores que puedan proporcionármela.

H. W. Becker, en un trabajo sobre los números de Bell publicado en 1941, presentó lo que llamaba "un interesante teorema". Se dice que una estrofa está completamente rimada si cada uno de sus versos rima al menos con algún otro, e incompletamente rimada si hay al menos un verso que no rima con ninguno de los restantes. Entonces, el número de estrofas completamente rimadas posibles con n versos es siempre igual al número de posibles estrofas incompletamente rimadas con $n-1$ versos.

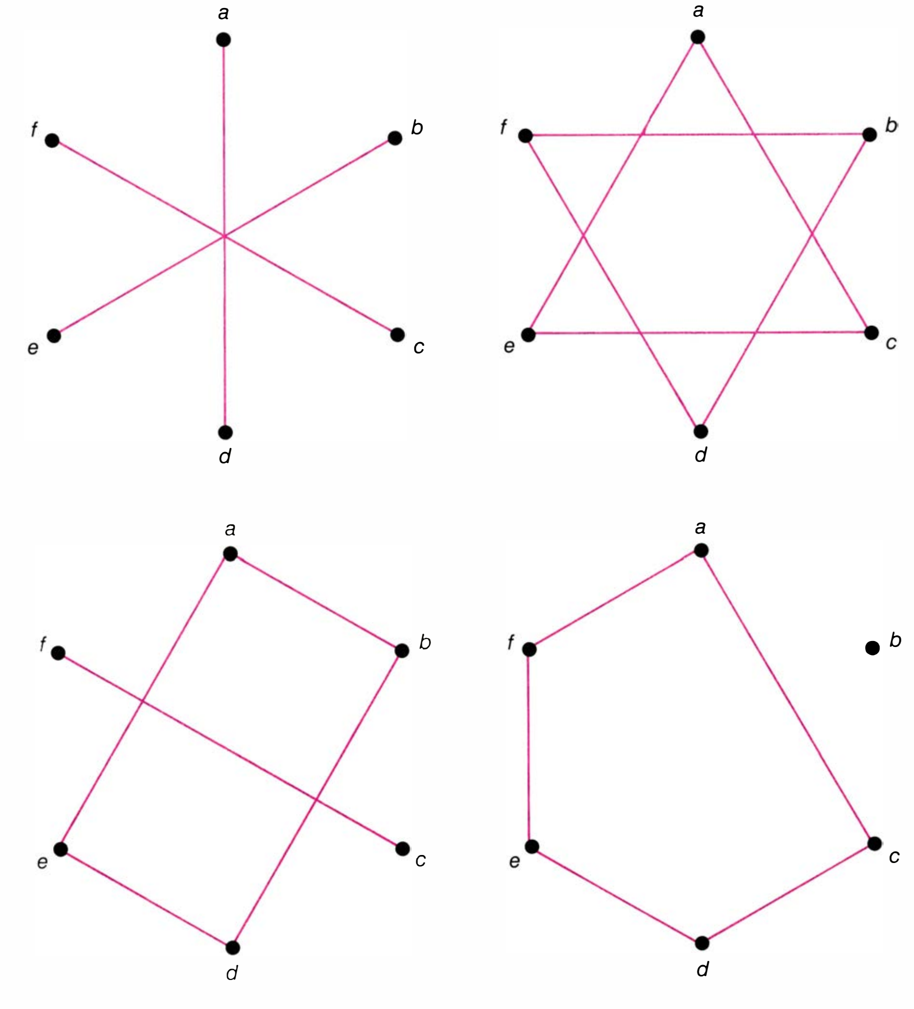
Henry W. Gould, especialista en teo-

ría de números de la Universidad de West Virginia, descubrió que los japoneses tenían ya en el año 1000 d.C. un método muy interesante de representar los esquemas de rimado. La ilustración de la página anterior presenta los 52 diagramas correspondientes a estrofas de cinco versos.

Las líneas verticales corresponden a los versos de la estrofa, mientras que las horizontales unen entre sí los versos que riman. Gould describió por primera vez los diagramas en 1976, en su *Research Bibliography of Two Special Number Sequences,* valioso catálogo de 175 referencias sobre los números de Bell y de 445 sobre los números de Catalán. Los lectores interesados en el tema pueden obtener un ejemplar enviando 3 dólares a Gould, 1239 College Avenue, Morgantown, W.Va. 26505.

La primera aparición de este método de representación del que se tiene noticia se encuentra en *The Tale of Genji,* famosa novela japonesa escrita por Lady Murasaki, quien vivió aproximadamente desde 978 d.C. hasta 1031. Todos los capítulos, excepto el primero y el último, de su libro, que consta de 54 capítulos, están encabezados por uno de los 52 diagramas para estrofas de cinco versos. Las líneas verticales son barritas de incienso, cada una de las cuales puede ser de un color a elegir entre cinco diferentes.

Las líneas horizontales unen barritas de igual color. En las primeras ediciones de este clásico japonés aparecían los diagramas en color; no así en las traducciones inglesas. Como hizo notar Joanne Growney en su tesis doctoral, en 1970, si se omiten todos los diagramas que contengan líneas que se intercepten, el número de diagramas restantes es el quinto número de Catalán, igual a 42, y este resultado sigue siendo verdadero en general para diagramas de Murasaki de *n* versos. Las razones que pudo tener Lady Murasaki para elegir precisamente este orden de representación de sus diagramas, escribe Gould, nos es tan desconocido como el fundamento, si es que hay alguno, de la ordenación de los 62 hexagramas de *I Ching*.



Un problema de Bell

En la poesía inglesa no abundan las quintillas, pero quizás una búsqueda diligente proporcione notables ejemplos de cada uno de los 52 ejemplos anteriores. Por ejemplo, el poema de Shelley "To a Skylark" está escrito en estrofas de cinco versos que se ajustan al quinto diagrama de la quinta fila. El cuarto diagrama de la segunda fila muestra el esquema de la conocida estrofa de Emily Dickinson:

To make a prairie it takes
  a clover and one bee,
One clover, and a bee,
And revery.
The revery alone will do,
  if bees are few.

[Hacer una pradera exigiría
un trébol, una abeja, y bastaría
ya... con fantasía.
Aunque la fantasía, si tú la dejas,
se bastará ella sola, si no hay abejas.]

He aquí una preciosa estrofa de "A Dead Harvest" de Alice Meynell, que se ajusta al modelo del cuarto diagrama de la cuarta fila:

Along the graceless grass of town
They rake the rows of red and brown,
Dead leaves, unlike the rows of hay
Delicate, touched with gold and gray
Raked long ago and far away.

[Por las feas avenidas ciudadanas
rastrillan los macizos, pardos y granas.
Hojas muertas, tocadas de oro y gris,
delicadas, que ya no sentís,
que el rastrillo arrastrará de aquí.]

El limerick es una estrofa de cinco versos con un esquema de rima ajustado al cuarto diagrama de la fila seis. Un limerick poco ortodoxo, atribuido a W. S. Gilbert, tiene el esquema del primer diagrama:

There was an old man of Dundee
Who was stung on the arm by a wasp.
    When asked "Does it hurt?"
  He replied "No, it doesn't.
  I'm so glad that it wasn't a hornet."

[Había un viejo de Dundee
que fue picado en un brazo por una
avispa
    Le preguntaron "¿Duele mucho?"
    Y contestó "Apenas nada,
¡cuánto me alegro de que no fuera
  un avispón!"]

Los números de Bell se utilizan de diversas formas en teoría de grafos. Tomemos el siguiente problema: se señalan seis puntos en una circunferencia, como si se tratase de los vértices de un hexágono invisible, y se marcan con las letras *a, ..., f.* Un punto aislado se considera que es un polígono convexo degenerado y que dos puntos, unidos por una línea recta son un polígono convexo degenerado de dos vértices. Usando un lápiz se conectan los puntos arbitrariamente, formando polígonos convexos disjuntos de uno, dos, tres, cuatro, cinco o seis vértices. (Disjuntos significa que no hay dos polígonos que posean un vértice común.) Los lados de un mismo polígono no pueden cortarse, pues entonces el polígono no sería convexo; sin embargo, no hay inconveniente en que se corten los lados de polígonos diferentes. Si se quiere puede no dibujarse ninguna línea, con lo que el modelo resultante estará formado por seis polígonos de un único vértice. También pueden conectarse los seis puntos y hacer un único hexágono. O bien, producirse una combinación cualquiera de polígonos, con tal de que sean convexos y disjuntos.

La ilustración de la página anterior muestra cuatro posibles modelos. ¿Cuántas configuraciones diferentes existen en total? Si ha comprendido el análisis de los números de Bell, tal cuestión no debería crear ninguna dificultad. Responderé a esta cuestión el próximo mes, dando una solución general válida para *n* puntos conectados por polígonos convexos disjuntos, de uno a *n* vértices.

En el artículo de abril afirmé incorrectamente que la ilustración inferior de la página 104 era proyección de un símplice de dimensión cuatro. Varios lectores escribieron haciéndome notar que la figura mostrada no es un símplice sino una sección de un complicado poliedro llamado poliedro hessiano. Puede verse una descripción de los poliedros hessianos en el maravilloso libro de H. S. M. Coxeter titulado *Regular Complex Polytopes* (Cambridge University Press, 1974, páginas 119-124).

William Funkenbusch escribió comunicándome que los dados de Sicherman solamente podrían utilizarse en la mesa de dados de un casino para realizar el juego en sí, pero que no se podrían utilizar para ciertos envites lícitos que requieren sumar las puntuaciones de los dados ordinarios. Por ejemplo, dos envites tales como sumar 4 o sumar 10 son, evidentemente, imposibles con dados de Sicherman.

Taller y laboratorio

Lo que los penachos de humos nos dicen sobre las características de la atmósfera

Jearl Walker

La observación de los penachos de humo en el cielo puede ser, a la vez, interesante e instructiva. Es mucho lo que se puede aprender en tal espectáculo sobre cómo varía la temperatura de la atmósfera con la altura. Los penachos de humo adoptan gran variedad de formas, determinadas principalmente por la diferencia de temperatura existente entre el aire de la superficie y el que se encuentra un poco más arriba que la chimenea.

Hay dos tipos básicos de penachos de humo: el de salida a chorro ("momentum jet"), en el cual el impulso hacia arriba del penacho se debe a su propia cantidad de movimiento inicial, y el penacho flotante; éste sube o baja según sea la relación entre su temperatura y la del aire. En los días con viento en calma, ambos tipos de penacho se elevan tomando la forma de un cono vertical invertido, que se va agrandando y cuyo vértice se halla muy cerca de la boca de la chimenea (suponiendo, en el caso del penacho flotante, que el gas que lo forma esté suficientemente caliente). En estos conos, tanto en uno como en otro tipo de penacho, la velocidad es mayor en el eje central del cono y va disminuyendo en dirección hacia el exterior.

Podemos establecer una diferencia entre ambos tipos de penachos: el de salida a chorro puro tiene un ángulo en el vértice del cono de aproximadamente 24 grados, mientras que el flotante puro tiene un ángulo en el vértice de 18 grados, también aproximadamente. No he podido dar con una explicación sencilla y a la vez convincente de tal divergencia. Las cifras anteriores se han deducido a partir de la observación de muchos penachos. Fotografiando un penacho vertical en un día de calma y luego midiendo el ángulo del cono en la fotografía, se puede determinar qué tipo de penacho es el que hemos observado. Si el ángulo es menor de 18 grados, ello se debe a que hay una corriente vertical ascendente, próxima a la chimenea, que hace que el penacho ascienda más rápidamente y, por tanto, hace decrecer el ángulo de su cono.

Un penacho es flotante (y por tanto asciende) si los gases que lo forman están más calientes que el aire a su misma altura, aunque también intervienen otros efectos físicos. Dado que la presión atmosférica decrece con la altura, toda porción de gas del penacho que ascienda debe expandirse al subir. La energía necesaria para expandirse la toma de la energía de las moléculas que lo forman, y no de ninguna fuente exterior. Esta energía interna es la energía cinética de las moléculas en su movimiento arbitrario.

La temperatura de un gas nos da una medida de su energía interna; mientras mayor sea su temperatura, mayor será la energía cinética de sus moléculas. Si parte de dicha energía se ha consumido en aumentar el volumen del gas, su temperatura habrá descendido necesariamente. Dicho cambio de temperatura, que no implica un intercambio de calor con el medio que rodea al gas, recibe el nombre de cambio adiabático. Por tanto, toda porción de gas, al ascender en un penacho de una chimenea, se expande enfriándose adiabáticamente. De igual manera, si se obliga a descender a una porción de gas de un penacho, ésta se contraerá y se calentará adiabáticamente conforme el aire que la rodea actúe sobre ella haciéndola bajar.

En un gas que sale por una chimenea, la continuación o no de su ascensión dependerá del cambio que experimente la temperatura del aire ambiente en relación con la altura. Por ejemplo, si un penacho al salir se enfría, pero aún así sigue siendo más caliente que el aire ambiente, continuará ascendiendo de forma acelerada. Por otra parte, si el gas llega a ponerse más frío, en su ascensión, que el aire que lo rodea, detendrá su ascenso

Un penacho de humo tipo vórtice o remolino

y comenzará a bajar. Si la temperatura se iguala, el penacho subirá a velocidad constante.

Cuando un gas se mueve inicialmente hacia abajo, se calienta adiabáticamente, pero su temperatura puede ser superior, inferior o igual que la temperatura del aire ambiente en cada momento, dependiendo de cómo cambie la temperatura del aire en relación con la altura. En el momento en que la temperatura del gas sea mayor que la temperatura del aire, detendrá su descenso. Mientras sea menor, el descenso proseguirá incluso de forma acelerada. Si la temperatura es la misma, continuará descendiendo pero a velocidad constante. Muchas de las formas características de los penachos que se pueden observar dependen de esta interrelación entre la ascensionalidad y las relaciones de enfriamiento o calentamiento en consecuencia de cómo varíe la temperatura de la atmósfera en relación con la altura.

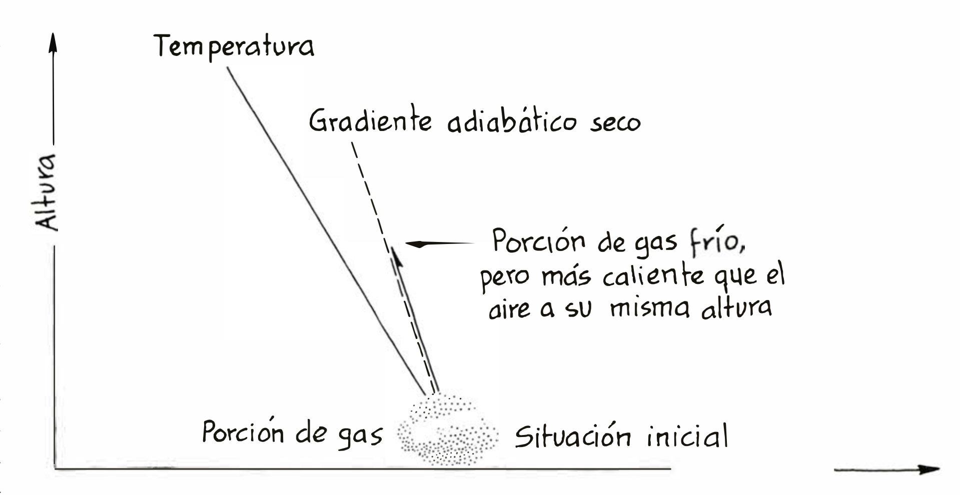
Toda porción de gas de un penacho de humo cambia su temperatura adiabáticamente –lo mismo al subir que al bajar– a razón de un grado Celsius por cada 100 metros de variación de la altura, aproximadamente. Esta razón se llama gradiente adiabático seco; la palabra "seco" indica que no se produce condensación. Dependiendo de las condiciones atmosféricas y de las fuentes de calor de los alrededores, la temperatura del aire puede descender con la altura más o menos rápidamente, o incluso en la misma proporción en que lo hace el gradiente adiabático seco de temperatura. Cada uno de los tipos anteriores de variación es un ejemplo de un gradiente negativo para el aire, debido a que su temperatura, en cada uno de los casos, decrece con la altura. Sin embargo, la temperatura del aire puede llegar a crecer con la altura; es lo que se llama inversión térmica. El aire tiene entonces un gradiente positivo.

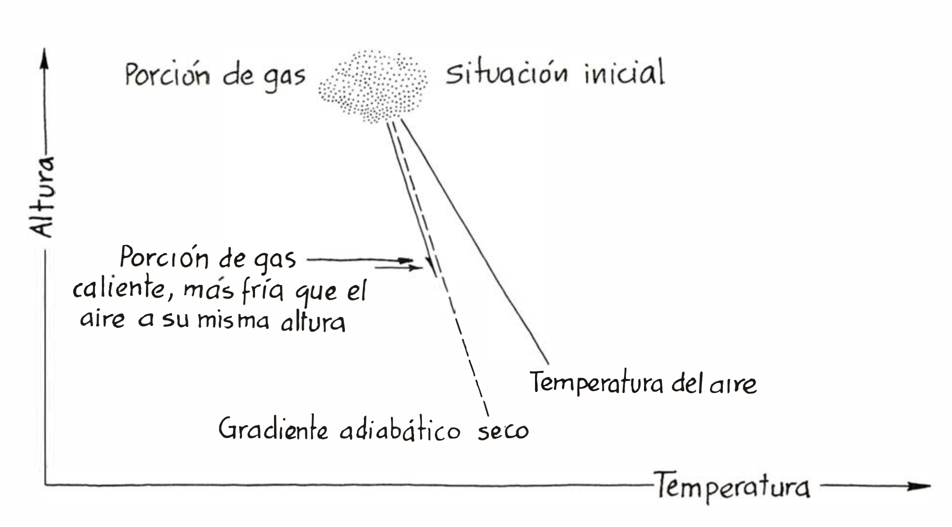
Si los gases están húmedos, el gradiente puede cambiar. La razón es que, conforme el gas se enfría, su humedad relativa se hace mayor; una vez que se alcanza el 100 por ciento de humedad, cualquier enfriamiento posterior provoca la condensación en forma de pequeñas gotitas de una parte del vapor de agua existente en el gas. El cambio de estado, de pasar de vapor a líquido, libera energía y evita que esta porción de gas se enfríe tan rápidamente como lo hubiese hecho de no existir el efecto de condensación. En otras palabras, parte de la energía necesaria para la expansión del gas se toma ahora del cambio de estado del agua, y, por tanto, se gasta menos cantidad de la energía interna del gas. La velocidad media a la cual una porción ascendente de gas cambia su temperatura con la altura se llama gradiente adiabático saturado. Cuando un gas húmedo comienza a ascender, se va enfriando de acuerdo con su gradiente adiabático seco, hasta que comienza a producirse la condensación; a partir de entonces, seguirá enfriándose, pero ya según el gradiente adiabático saturado. Esta variación en el ritmo o velocidad de enfriamiento afecta a los penachos húmedos del tipo de los que emiten las centrales térmicas; en ellos, los gases no se enfrían tan rápidamente mientras asciende como lo hacen en los penachos secos.

Cabe la posibilidad de que un penacho de los de salida a chorro, al expansionarse un tanto, se comporte básicamente como un penacho flotante. Existe otro tipo de penachos que se forman en los días de calma, y que a mí me resultan extraños. Son los que tienen una velocidad inicial de los gases muy pequeña, lo cual puede ocurrir cuando en las chimeneas caseras existe una abertura demasiado grande; por tanto, el penacho no se expande inicialmente como un cono sino que se contrae hasta que la velocidad ascensional se incrementa suficientemente y, a partir de entonces, se expansiona como un penacho flotante normal.

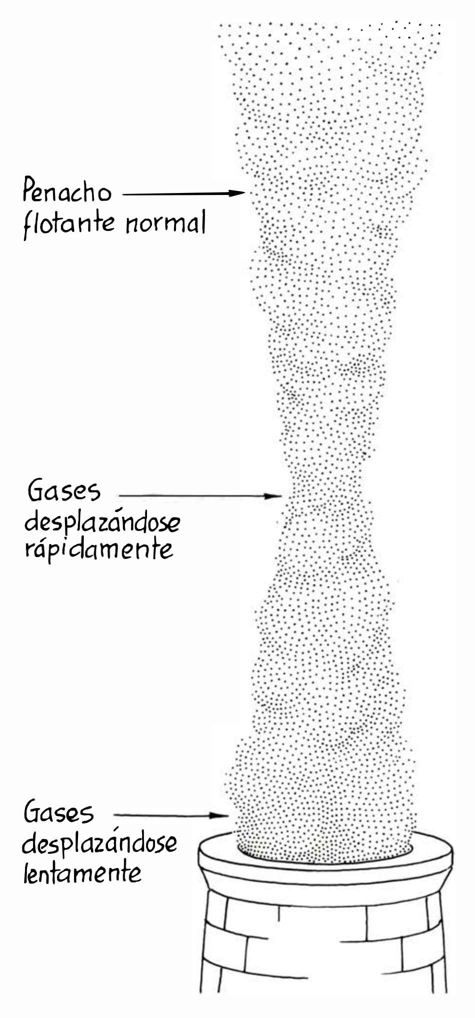
Se puede ver un ejemplo similar, aunque invertido, de este estrangulamiento en el fregadero de la cocina. La corriente de agua, al salir del grifo, cae, aumenta su velocidad y se estrecha. Este estrechamiento no se debe a la presión atmosférica, ya que dicha presión permanece constante a todo lo largo de la caída del agua.

Para ponerlo más claro imaginemos una sección transversal de la corriente de agua cerca de la salida del grifo y otra un poco más abajo. Cada segundo pasa una determinada cantidad de agua a través de la sección superior. La misma cantidad deberá pasar también por la sección inferior en el mismo tiempo; de lo contrario, tendríamos que, por arte de magia, cierta cantidad de agua aparecería o desaparecería entre los dos niveles. Puesto que la velocidad a través de la sección inferior es mayor, el agua necesita menor sección de paso para que





Comportamiento de una porción de gas al ascender (**arriba**) *o al descender* (**abajo**)



Un penacho flotante y convergente

fluya la misma cantidad cada segundo. La corriente, por tanto, se estrecha.

Los gases calientes al salir por una chimenea con poca velocidad no pueden expandirse de la misma forma que lo haría un penacho flotante; y ello porque la forma típica del cono requiere una determinada relación entre la velocidad inicial y la flotabilidad. Como consecuencia, los gases se elevan primero como un chorro, aceleran luego su ascenso merced a su flotabilidad y contraen su anchura conforme aumenta su velocidad. Cuando hayan alcanzado una velocidad suficiente como para conseguir la relación requerida, el penacho se expandirá formando un cono. En los días con viento en calma puede ser divertida la búsqueda y observación de penachos convergentes en las chimeneas de las casas o en las fogatas.

Para ver los penachos flotantes más bonitos hay que esperar los días de viento suave y constante, que inclina o dobla los penachos por la parte superior. Cuando han terminado de subir, descienden, se dispersan hacia arriba y hacia abajo, adoptan una forma alargada o forman bucles, según varíe la temperatura del aire de la zona en relación con la altura y según la fuerza del viento en dicha zona. La ilustración de la página siguiente muestra una amplia gama de penachos típicos junto con la variación de la temperatura del aire según la altura que se requiere para cada uno de ellos. El gradiente adiabático seco de temperatura también aparece allí, pero esa línea se puede mover hacia la derecha o hacia la izquierda. La gráfica indica la relación entre el cambio de temperatura y el cambio de altura para una porción de gas de un penacho. Es necesario conocer la temperatura del gas en algún punto determinado (por ejemplo, la parte alta de la chimenea) para que esta línea pueda representar la temperatura del gas en función de la altura. La gráfica entonces habría que moverla a la derecha o a la izquierda para que diera justo esa temperatura a esa altura.

Si la temperatura del aire decrece con la altura algo menos rápidamente de como lo hace el gradiente adiabático seco de temperatura del gas, se limita la ascensión al encorvamiento del penacho por su parte superior, y desciende. Un ascenso enfriaría adiabáticamente el gas con la consecuencia inmediata de que dicho ascenso se detendría. Un descenso calentaría el gas adiabáticamente, con lo que también el descenso se detendría. Aunque el perfil de temperatura de la atmósfera esté suficientemente cerca del gradiente adiabático, el penacho se va expandiendo gradualmente produciendo la llamada forma "de cono".

Por el contrario, cuando la temperatura de la atmósfera decrece mucho más rápidamente con la altura que el gradiente adiabático seco (a dicha forma de variación se le llama gradiente superadiabático), el calor de la chimenea o del suelo puede provocar grandes remolinos térmicos, suficientes como para atrapar el penacho y hacerlo girar y formar ondas hacia arriba y hacia abajo. Estas circunstancias ocurren en días cálidos, cuando el suelo absorbe rápidamente una cantidad de calor solar antes de que el aire circulante pueda llevárselo. Los remolinos nunca se forman cuando hay viento fuerte, que enfría el suelo, o éste se halla cubierto de nieve.

Si la temperatura del aire tiene un gradiente positivo, es decir, si dicha temperatura se incrementa con la altura, formando una inversión térmica, el penacho se limita a ser una fina corriente de gas que se abre en abanico, cuyo vértice estaría en la chimenea. Cuando una parte de dicho penacho intenta subir, se enfría adiabáticamente poniéndose bastante más frío que el aire ambiente, por lo que su ascenso se frena de inmediato. De modo similar, si una parte del penacho intenta descender, se calienta adiabáticamente poniéndose más caliente que el aire ambiente y, por tanto, frenando su descenso de forma inmediata.

Cuando se produce una inversión térmica por debajo del vértice de la chimenea que impide los movimientos descendentes del humo, la figura que éste forma se la conoce por penacho en elevación. Este tipo de variación de temperatura suele darse al atardecer, cuando el suelo y la parte baja de la atmósfera han comenzado a enfriarse a medida que el sol se va escondiendo, mientras que la parte de la atmósfera próxima a la boca de la chimenea está aún relativamente caliente.

Fumigación es el término que se emplea para el fenómeno que ocurre cuando un gradiente positivo y negativo se dan a la vez con una inversión sobre la región donde el gradiente es negativo. La inversión impide que el penacho se difunda hacia arriba; al calentar entonces los humos el suelo, se provoca la aparición de remolinos que arrastran los humos hasta el mismo nivel de la superficie. Esta situación ocurre con frecuencia por la mañana, cuando el suelo y la parte baja de la atmósfera comienza a calentarse.

Un perfil de temperatura similar, pero cuyo gradiente sea menos empinado cerca del suelo, da lugar a un efecto sifón entre el nivel en el que esté la inversión y el suelo. El penacho se difunde hacia el suelo sin que sea necesaria la existencia de una gran turbulencia provocada por la mayor temperatura del suelo.

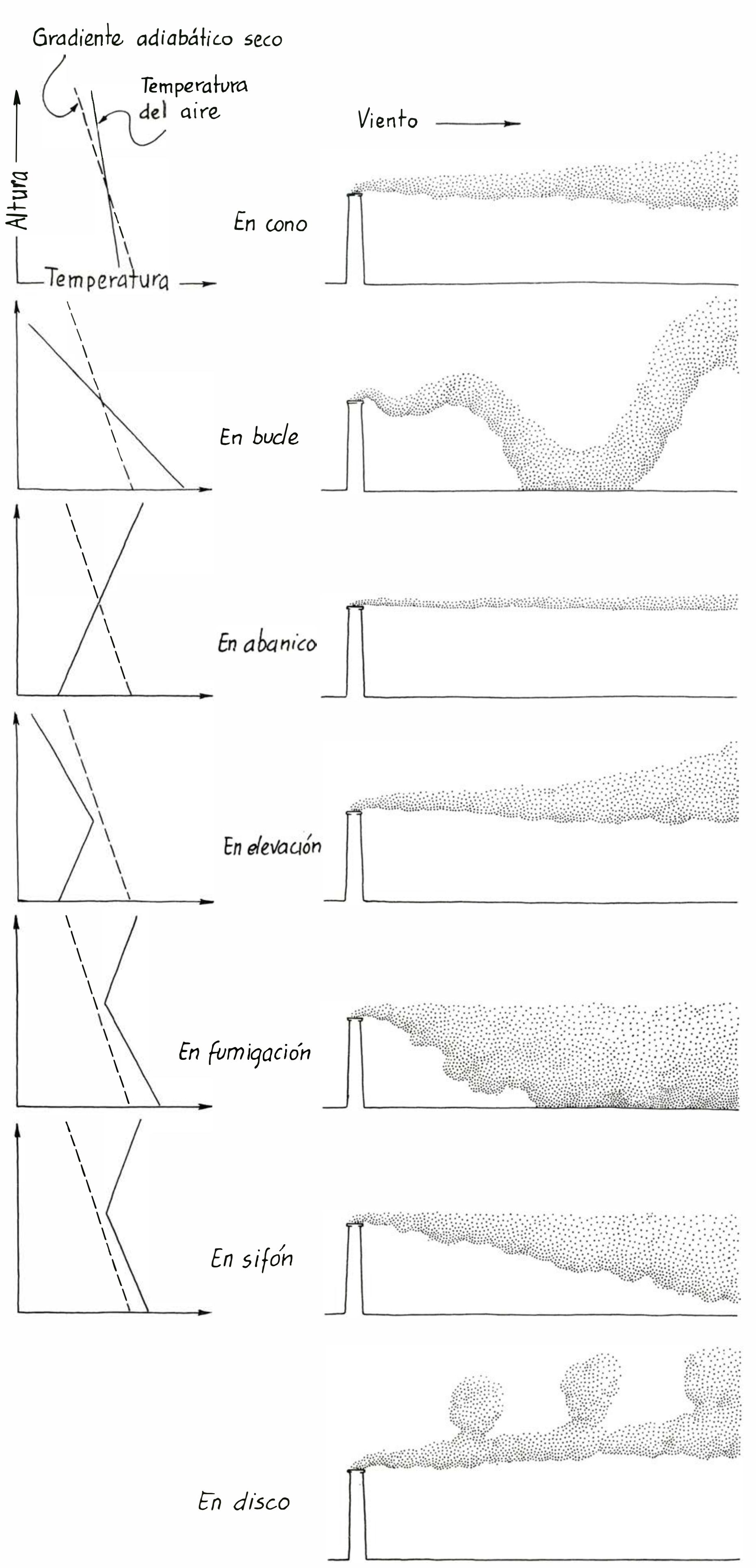
Mas si el penacho se produce en forma de cono o en elevación, puede ser lo suficientemente caliente como para lanzar hacia arriba chorros de aire caliente, dando lugar a que asciendan burbujas de aire caliente con forma de buñuelos, con un intenso movimiento ascendente por el centro y otro movimiento, más débil, descendente, por la parte de fuera. Tales corrientes, que se dan en las zonas de la atmósfera por encima de las fábricas, son aprovechadas por los pilotos para ganar altura.

La forma del penacho en cada uno de estos casos fundamentales puede modificarse si el humo abandona la chimenea con una gran cantidad de movimiento. Por ejemplo, si en las situaciones de fumigación o de sifón la capa con inversión térmica es relativamente delgada, un penacho con una cantidad de movimiento inicial suficiente puede penetrar en dicha capa en lugar de permanecer

confinado debajo de ella. En algunas localidades en donde se registra con frecuencia o de forma persistente la inversión térmica, suelen diseñarse sus chimeneas de suerte que, en su salida, los gases reciban una cantidad de movimiento suficiente como para que penetren en la capa de inversión antes de que el penacho se ponga horizontal. Así se reduce considerablemente la polución a nivel del suelo.

Las formas que adoptan los penachos, según sean húmedos o secos, divergen un tanto debido a que el gradiente adiabático es menos pronunciado una vez que el agua contenida en los gases comienza a condensarse. Algunos penachos, los que se producen en las centrales térmicas por ejemplo, son visibles gracias al agua que se condensa casi inmediatamente después de salir, produciendo una estela blanca. Sin embargo, la mezcla de los humos con el aire ambiente provoca rápidamente la evaporación, de ahí que la parte visible del penacho sea pequeña. En algunas ocasiones los humos o gases son lavados antes de salir por la chimenea, haciéndolos pasar por unas duchas que rocían agua con el fin de quitarles polvo o sustancias contaminantes. En esos casos, los gases, al salir por la chimenea, llevan realmente agua en estado líquido, de suerte que cuando el agua se evapora, al mezclarse los gases con el aire ambiente, el penacho se enfría en virtud del gasto de energía que ha sufrido en la evaporación. (Cuando nos secamos la piel soplando sobre ella, sentimos frío debido a que es el propio cuerpo el que suministra la energía necesaria para evaporar el agua haciéndola pasar del estado líquido a vapor.) Como resultado de este rápido enfriamiento, el penacho caerá, llegando al suelo probablemente muy cerca de la chimenea.

Si se observa y se va fotografiando el penacho de una chimenea a lo largo de un día completo, se verá cómo cambia a medida que el suelo se vaya calentando o enfriando. Al amanecer, la tierra y la parte baja de la atmósfera están frías; ello permite la formación de penachos en forma de abanico o en elevación. Conforme el suelo va absorbiendo energía solar, tanto éste como la capa inferior del aire se va calentando, y comienzan a formarse turbulencias. Esta cálida y turbulenta capa de aire se va haciendo cada vez más ancha, desarrollando las formas de fumigación o de sifón. Ambos efectos hacen que el penacho descienda hasta el nivel del suelo. Más tarde, cuando se ha calentado ya el aire de la zona superior, el penacho puede desarrollarse en forma de cono, pueden salir de él una especie de chorros en disco de aire o (si el calen-



Tipos de penacho creados según la variación de la temperatura del aire con la altura

tamiento del suelo ha sido muy fuerte) puede adquirir forma serpenteada, en bucle. Después del mediodía, cuando la intensidad de los rayos del sol disminuye, y tanto el suelo como las capas bajas del aire comienzan a enfriarse por radiación, todos los penachos pueden constituir conos. Al anochecer, en la zona baja de la atmósfera se desarrolla una inversión debido al enfriamiento por radiación del suelo; el penacho comienza a adquirir forma de abanico o en elevación de nuevo.

Cabe la posibilidad de que alguien observe simultáneamente varias chimeneas de distintas alturas. En Cleveland puedo ver al menos una docena de chimeneas en un valle poco profundo, y poniéndome en un punto desde donde se domine el valle puedo comparar los tipos de comportamiento de los penachos de humo. Con una distribución apropiada de la temperatura, los penachos de humo emitidos por las chimeneas más altas pueden diferir de los emitidos por las más bajas debido a que los dos conjuntos de humo están en partes diferentes del perfil de temperatura atmosférica. Por ejemplo, las chimeneas altas pueden emitir humo a una zona con un gradiente negativo y mostrar un penacho en elevación, aun cuando las chimeneas más bajas emiten a la inversión y se extienden en forma de abanico. Si se observa un número suficiente de chimeneas se puede estimar dónde se produce el cambio de gradiente negativo a positivo. En días en que hay mucha inversión a baja altura se pueden ver algunos penachos de humo que "pinchan" la inversión mientras que otros se mantienen debajo.

Un penacho de humo puede adherirse a su chimenea. A veces la razón de esto es que el viento forma torbellinos y empuja el humo hacia abajo, a lo largo de la chimenea. Otras veces el descenso del humo puede estar motivado por la situación de la chimenea con relación a los edificios y columnas circundantes. Por ejemplo, si la chimenea está en el lado de sotavento de un gran obstáculo, el viento es obligado a ascender por mor del obstáculo y desciende después por el otro lado, haciendo que el humo también descienda. Si, por contra, la chimenea está en el lado de barlovento, es probable que el humo se eleve empujado por el flujo ascendente del viento.

Si la abertura de la chimenea es demasiado grande para la cantidad de gas que expulsa, y si esta expulsión es lenta, el humo saldrá a bocanadas. Unas veces la boca emitirá gas y humo de forma normal, otras el aire frío del exterior penetrará por la chimenea impidiendo la expulsión. Llegará un momento en que los gases calientes se unan de nuevo y consigan salir otra vez por la chimenea, produciendo otra bocanada de humo, hasta que se vuelva a repetir el proceso.

Si la chimenea expulsa un penacho flotante que no sufre mucha mezcla por turbulencia, éste se bifurcará casi inmediatamente. La velocidad es mayor en el centro del humo expulsado que en el exterior, lo que obliga al humo a circular formando dos torbellinos, de dirección ascendente en el medio y descendente en el exterior. Los dos torbellinos resultantes dividen el penacho de humo. Esto lo puede comprobar cualquier observador si tiene la oportunidad de acercarse a una chimenea que expulse un penacho flotante en un día de viento tranquilo y sin turbulencias. Deberá colocarse debajo del humo y levantar la vista para fijarse en su centro. Es probable que el movimiento envolvente sea bastante claro mientras se mira al humo saliendo y moviéndose hacia abajo. Y hasta puede verse el firmamento entre las dos corrientes.

Para hacer un estudio serio del comportamiento del humo, debe intentarse relacionar la mezcla descendente del humo con la forma en que éste se mueve y con las condiciones del suelo situado debajo de éste. Por ejemplo, una chimenea puede emitir un humo que toma la forma de un abanico durante todo el tiempo que permanece sobre la zona del suelo que está relativamente fría y tiene pocos obstáculos para crear una turbulencia a gran escala. Si el penacho con forma de abanico pasa sobre una masa de agua, es probable que el agua y el aire que están entre ambos se hallen más calientes que la tierra. La turbulencia asociada con el agua y el aire más calientes produce fumigación, lo cual hace descender al humo hasta el nivel del agua. He visto chimeneas en Cleveland que emiten penachos finos de humo marrón que pasaban sobre el lago Erie, en donde se producía la fumigación y el material marrón descendía y se mezclaba al nivel del agua.

La fumigación puede desarrollarse también a partir de penachos en abanico si éstos pasan sobre fuentes de calor artificiales, como una ciudad por ejemplo. Allí el calor mezcla el aire hasta un nivel de unas tres veces la altura media de los edificios. Si el observador pudiera encontrar una chimenea humeante que fuera visible desde muy lejos, se podría ver el paso del humo desde la forma de abanico a la de fumigación, a medida que éste atraviesa una fuente de calor natural o artificial.

Los accidentes del terreno pueden afectar de otras formas al humo. Un valle estrecho puede ayudar a mantenerlo cerca del suelo si se desarrolla una inversión térmica sobre el aire frío atrapado en el valle. Cualquier penacho de humo que se cree debajo de la inversión

Altura
Capa superior con inversión ténmica durante el día
Máxima altura de calentamiento durante el día
Noche
Día
Temperatura

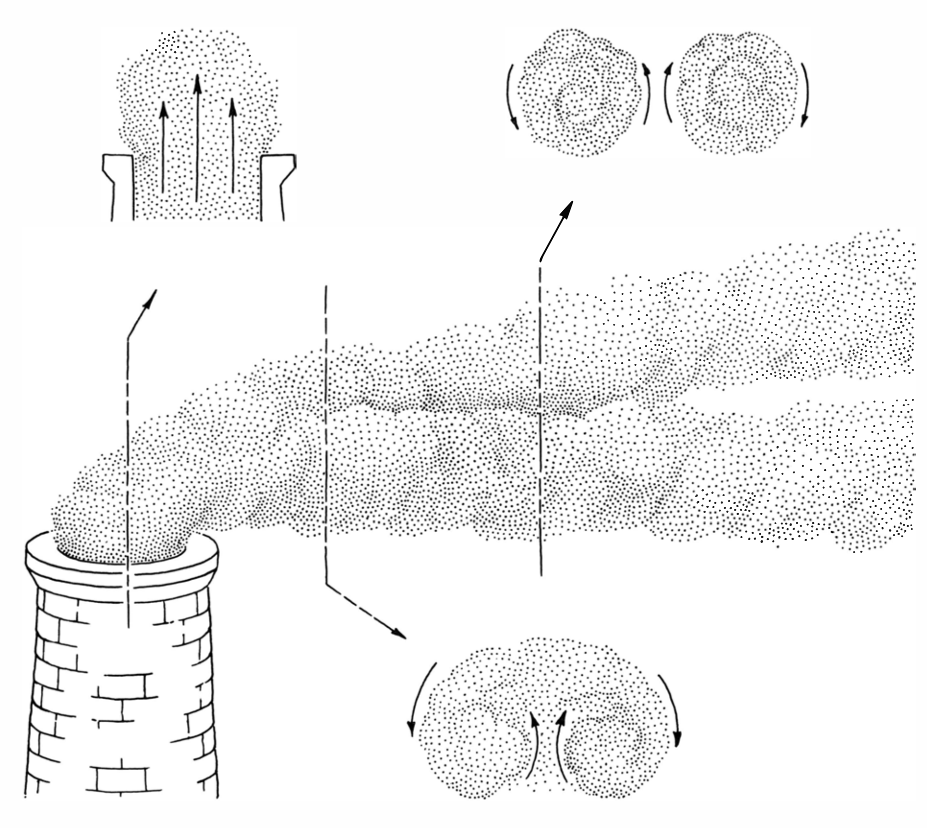
Variación diaria de la temperatura del aire con respecto a la altura

no podrá penetrar la capa de ésta y escapar a no ser que tuviera una cantidad de movimiento a la salida lo suficientemente alta para conseguir la penetración mientras se estuviera aún comportando como un penacho de salida a chorro.

Esta retención de humos nocivos en algunos valles ha ocasionado varios desastres. En el valle del Mosa, en Bélgica, se desarrolló una inversión en diciembre de 1930, reteniendo en la cuenca las emisiones de las acerías y de las plantas químicas. Como consecuencia de esto murieron 60 personas en tres días, la mayoría por fallo respiratorio; se contaron por miles las que contrajeron enfermedades graves. Es difícil creer que las autoridades, en lugar de interrumpir las emisiones, achacaran el incremento de diez veces en el número de muertes a una extraña y desconocida enfermedad, llegando incluso a sostener un biólogo que la causa era una plaga. Algún tiempo más tarde, se culpó a la polución industrial, pero se hizo muy poco para atajarla. En septiembre de 1972, el valle volvió a sufrir una inversión térmica que condujo de nuevo a la retención de la polución y fue causa de muchas enfermedades, aunque esta vez no mortales.

La famosa niebla que caracterizó a Londres en otros tiempos se debía al abundante uso del carbón en las fábricas y en las casas, que vertían diariamente toneladas de anhídrido sulfuroso al aire. Al condensarse el agua en las partículas, se formaba una niebla espesa y marrón (la llamada "puré de guisantes"). En diciembre de 1952 se lanzaron a la atmósfera ambiente unas 2000 toneladas de anhídrido sulfuroso cada día, y se produjo una inversión térmica sobre el aire más frío del suelo que mantuvo retenido el humo de las casas y fábricas por debajo de ellas. Debido a la ausencia de viento que lo dispersara, el anhídrido sulfuroso se concentró sobre la ciudad durante cuatro días y cuatro noches, tornándose el cielo primero amarillo, después marrón y, finalmente, negro. A consecuencia directa de la polución murieron unas 4000 personas y otras 8000 murieron después por complicaciones respiratorias causadas por la fuerte inversión.

Si un observador dispone salir a observar el humo de las chimeneas, pero el día resulta ser demasiado lluvioso, puede quedarse en casa y examinar el humo procedente de un cigarrillo. Cerca de la punta encendida, la corriente de humo es uniforme y delgada, unos centímetros más arriba se rompe y forma remolinos para acabar dispersándose. Para distinguir esta transición hay que estar en una habitación al abrigo de corrientes de aire o en un tubo de vidrio grande abierto por ambos extremos. Cuando ascienden los gases procedentes de la punta encendida tienen mucha fuerza, porque están más calientes que el aire de la habitación. Sin embargo, dado que la velocidad inicial es relativamente baja, los gases calientes suben formando una corriente fina. Después de algunos centímetros, quizá de unos 30, la aceleración hacia arriba ha incrementado la velocidad hasta el punto en que la corriente se hace inestable y empieza a romperse y a formar remolinos.


Un penacho dividiéndose en dos remolinos

En el agua que circula por las tuberías sucede una transición similar, que pasa de corriente laminar a turbulencia. Si el agua circula despacio (el significado de "despacio" depende del radio de la tubería y de la densidad y viscosidad del agua) se mueve con suavidad. Cada pequeña porción de agua fluye a lo largo de una línea paralela a la pared de la tubería. Pero a una determinada velocidad ya no fluye con la misma suavidad, sino que se mueve de manera inestable, se forman remolinos y cada pequeña cantidad de agua sigue una ruta irregular a lo largo de la tubería. Estos remolinos se pueden hacer visibles en una tubería transparente echando colorante al agua. En la corriente de humo producida por un cigarrillo, las partículas de éste nos permiten ver el curso seguido en su evolución.

Si el cigarrillo no tiene filtro, se puede notar una lenta corriente de humo que emerge del extremo no encendido. Esto es debido a que los gases que transportan el humo, después de haber pasado a lo largo del cigarrillo, se han enfriado, salen a la temperatura que reina en el ambiente, más o menos, y, por tanto, carecen de fuerza ascendente.

Los indios americanos usaban fogatas de humo para comunicarse entre grandes distancias. En el libro "Natural Aerodynamics", R. S. Scorer describe cómo los aborígenes australianos enviaban señales de humo a mayores distancias todavía. Mientras unos construían una fogata con grandes palos, otros echaban encima arbustos verdes. La altura de la fogata incrementaba la circulación del aire a través del fuego, se avivaba el fuego y los arbustos producían así un humo espeso. Con un esfuerzo coordinado de los trabajadores se puede llegar a hacer ascender una térmica llena de humo. Cuando se hizo esto con la inversión de temperaturas de la montaña o del atardecer, la térmica ascendía hasta que su temperatura era igual que el aire del ambiente, en cuyo punto el humo se difundía horizontalmente para dibujar una forma parecida a un champiñón gigante. Los trabajadores podían controlar la altura en la cual se realizaría la expansión horizontal, estimando cuidadosamente el grado de calor y de altura de la fogata. Si la ocasión era especialmente significativa, los trabajadores podían dibujar hasta seis champiñones sobre la fogata.

Libros

Del enigma de la evolución de los seres a la redomada picaresca por ordenador

Enrique Gadea y Philip Morrison

La evolucion de lo viviente, por Pierre P. Grassé. Traducción del francés por Ernesto Fernández y Arantza Plazaola. H. Blume, Ediciones; Madrid, 1977; 330 páginas. La edición original de esta obra se publicó en Francia con el título "L'évolution du vivant", en 1973, en la colección "Sciences d'aujourd'hui" (Edit. Albin Michel, París). Su autor es una figura del mayor prestigio dentro del campo de la biología; miembro actualmente de la Academia de Ciencias de París, ha sido profesor titular de Evolución de los Seres Organizados en la Facultad de Ciencias de la Sorbona durante más de treinta años. El profesor Grassé es tal vez el máximo representante de la zoología actual en su sentido más genuino. Su famoso "Traité de Zoologie" es un auténtico monumento de la zoología universal que marcará un hito en la ciencia de todos los tiempos. Pero, además, el profesor Grassé es un auténtico naturalista, un biólogo en su sentido integral, a la par que un gran viajero y un profundo humanista: es un hombre de ciencia que ha escapado a los excesos de la especialización y sabe concebir y hacer ver la biología en su sentido más natural y más amplio. El hombre y la evolución han llegado a preocuparle extraordinariamente, exponiendo sus concepciones en este campo en diversos trabajos y obras, siendo la que aquí se comenta la más elaborada y representativa. En ella, con una preparación exhaustiva para tratar de la evolución del mundo viviente, sin ideas preconcebidas y limpio de todo dogma, expone las líneas maestras de un nuevo enfoque de la misma, apoyándose en las formas fósiles y en las últimas conquistas de la biología molecular y de la genética.

Hoy en día se considera la evolución no como una hipótesis, sino como un hecho, sin cuya aceptación carece de sentido el mundo viviente. Sólo aquellos cuyas creencias dogmáticas actúan como una barrera mental frente a lo evidente –dice el profesor Grassé– pueden negarla o ponerla en duda. Sin embargo, cuando se trata de explicar su mecanismo, surgen las discusiones y el desacuerdo entre los biólogos, persistiendo todavía las ya viejas controversias entre las concepciones lamarckianas y darwinistas, en tanto que las pretensiones neodarwinistas se tambalean cada vez más ante la complejidad de los hechos. A este respecto el autor pone en evidencia que hasta ahora ninguna doctrina ha explicado verdaderamente el mecanismo evolutivo.

La evolución biológica tiene como fin el conocimiento y estudio del origen de los seres vivos, de sus transformaciones y de la especulación sobre su posible futuro. Ignoramos si los primeros seres organizados que aparecieron sobre el planeta fueron procariotas quimiótrofos, como se supone. Lo que sí podemos afirmar es que la vida no tiene lugar más que cuando determinadas estructuras, compuestas de grandes y complejas moléculas orgánicas de naturaleza particular, están ordenadas según un determinado modelo: la célula. La primera etapa franqueada por la vida después de su aparición fue consiguientemente la etapa unicelular. En el conjunto de la célula aislada la vida se ha mostrado sumamente creadora a lo largo de la evolución, dando lugar a los grandes tipos de esquizófitos, protófitos y protozoos. La segunda gran etapa de la evolución fue el paso al estado pluricelular y concretamente metacítico. Gracias a este último, el organismo se libera de las causas limitantes del tamaño de la unicelularidad, a la par que la especialización funcional de los territorios celulares conduce a la formación de los tejidos y posteriormente a los conjuntos funcionales superiores, como órganos, aparatos y sistemas. Ello permite a la materia viva entrar abiertamente en el camino de la gran expansión evolutiva, en principio ilimitada, de todo el mundo viviente.

No es necesario destacar aquí –como lo hace magistralmente Grassé en su obra– todos los niveles y etapas de este proceso evolutivo en los vegetales y animales. Sí que hay que decir, en cambio, que en el curso de esta génesis el máximo perfeccionamiento evolutivo lo han alcanzado los metazoos triblásticos en dos de sus grandes linajes: los artrópodos y los vertebrados. En estos últimos, el sistema nervioso ha marcado la pauta de su tipo estructural y de su progresiva y clásica evolución. En los mamíferos, el cerebro, al perfeccionarse hasta el máximo y notablemente al adquirir un neopalio (córtex), se convirtió en un órgano nuevo, que dotó a su poseedor de la facultad del pensamiento conceptual, de la libertad y de los principios morales. En este sentido el hombre hizo pasar la evolución a una esfera ya no sólo orgánica, sino también espiritual y social. La formación del hombre sólo fue posible en un medio social capaz de permitir la facultad de conservar las adquisiciones gracias a la tradición primero oral y luego escrita. Sin la evolución paralela de lo orgánico y lo social, el *Homo sapiens* no habría nacido nunca, sostiene el profesor Grassé.

En cuanto a los mecanismos que han obrado y obran en el proceso de la evolución y la biocinesis, donde radica el auténtico meollo de la cuestión, el autor busca la explicación basándose en los datos de la paleontología, a la que considera primordial fuente de información, ya que refleja la auténtica historia de los organismos en el planeta, junto con las aportaciones de la moderna genética y de la biología molecular. Con estos criterios se adentra en un nuevo camino y ofrece una concepción objetiva y prospectiva de la evolución y de sus mecanismos, esbozando una plausible teoría general de la misma, cuyos puntos cruciales se resumen a continuación.

La evolución de lo viviente depende esencialmente de procesos que se efectúan a nivel de infraestructuras y se desencadena por factores internos y externos. Es un fenómeno orientado, que no se origina sólo de variaciones hereditarias aleatorias. La evolución exige la adquisición, a lo largo de los tiempos y a medida que los organismos se complican, de novedades cuya información se inserta en las hélices del ADN bajo la forma de nuevos genes. Este fenómeno es profundamente diferente de la mutagénesis productora de alelos. La paleontología revela que los linajes de un tronco o filo común presentan la misma tendencia para organizarse en un mismo tipo estructural, aunque en grados diferentes. Como corolario, termina diciendo el profesor Grassé que los esfuerzos conjuntos de la paleontología y de la biología molecular (desembarazada ésta de sus dogmas) deberían conducirnos al mecanismo exacto de la evolución; eso sí, sin revelarnos las causas de su orientación y finalidad, dominio éste en que la biología se siente impotente y debe ceder, al menos por ahora, la palabra a la metafísica.

Los capítulos de la obra son realmente sugestivos, tanto por la claridad de ex-

posición, como por su contenido. Tras presentar el problema de la evolución en sus consideraciones generales, el autor trata de sus aspectos progresivos y regresivos y del nacimiento de los tipos de organización, todo ello con una documentación muy bien asimilada y meditada. Pasa luego a estudiar la evolución como fenómeno histórico discontinuo y el papel que el azar puede llegar a jugar en ello, enfocado con un sentido crítico encomiable. Resulta extraordinariamente interesante el capítulo sobre evolución y selección natural, así como el dedicado a la adaptación y a su papel en este proceso. Con fino sentido filosófico y científico a la vez, trata la cuestión de la evolución y la necesidad, distinguiéndola de la utilidad. Interesantísimo es también el capítulo dedicado a las actividades génicas en relación con la evolución. En el último capítulo expresa el autor sus puntos de vista sobre la nueva interpretación de los fenómenos evolutivos, constituyendo, en cierto modo, la síntesis de sus ideas sobre este polémico y sugestivo campo. Termina la obra con una conclusión que resume los puntos esenciales. Se completa con unos apéndices con documentos diversos, un glosario y un índice alfabético de materias y autores.

Esta obra del profesor Grassé es el fruto de la meditación de toda una vida que ha sabido asimilar, comprender e interpretar, como pocos, esa maravilla que es el mundo biológico a través del prisma de la evolución. Es un libro que se lee con fruición y que todo biólogo que en verdad se sienta naturalista debería conocer y meditar. Es una de esas obras que con el tiempo van adquiriendo cada vez más valor, por lo mucho del saber vivido y pensado que hay en ellas y que preludian convertirse en clásicas. (E. G.)

CRIME BY COMPUTER, por Donn B. Parker. Charles Scribner's Sons (10,95 dólares). COMPUTER SECURITY AND PROTECTION STRUCTURES, por Bruce J. Walker e Ian F. Blake. Dowden, Hutchinson and Ross, Inc. (14,95 dólares). Todo usuario de un teletipo terminal de computador espera "firmar" al empezar su labor. A menudo un simple golpecito en la tecla de retroceso del carro inicia el proceso. El terminal imprime la hora y fecha y un saludo de rigor, con algunas consabidas solicitudes de información, a las que uno responde como de costumbre con la identificación de crédito codificada y la contraseña que permite el acceso a determinado archivo privado. Sin mayor aviso previo, el sistema "se estrella": hay algún defecto, quizás una interrupción momentánea de energía eléctrica. Uno ha de firmar otra vez, siguiendo un procedimiento archisabido. Es una molestia nimia, que de modo alguno se nota. ¡Pero a veces hay una ponzoña! La primera vez que uno ha escrito la contraseña y los datos de identificación puede haber sido una pseudofirma. Ese programador posee ahora las contraseñas de uno, y algún día podrá explotarlas en provecho propio. Ya en 1970 un par de hábiles alumnos de bachillerato acumularon un centenar de contraseñas precisamente con esa trampa inconspicua.

Al alcance de cada terminal hay un mundo de historias personales, de bienes económicos y de transacciones efectivas, de secretos de personas y sociedades. En Estados Unidos hay 150.000 computadores en servicio, y su número se está triplicando cada cinco años. Parker, analista jefe de procesado de información del Stanford Research Institute, se ha convertido desde 1970 en estudiante a tiempo completo del problema del abuso del computador. Su libro es un resumen popular de lo que ha aprendido, de las estadísticas de las pérdidas, de las actitudes de la comunidad informática, un estudio de los perpetradores vistos como clase aparte, y, más que todo, el relato de una alegre docena de casos de notable variedad delictiva, casi todos ocurridos en los Estados Unidos. Walker y Blake ofrecen la otra cara de la moneda. La de ellos es una monografía sobria y compacta, llena de jerga, pero de argumentación no excesivamente técnica, que analiza y reseña la literatura sobre amenazas y contramedidas para el programador y el directivo: a ambos les resultará atrayente. Estos autores, ambos de la Universidad de Waterloo, Ontario, incluyen breves descripciones de las medidas de seguridad que se han puesto en práctica en una docena de diversos y conocidos sistemas de computador.

De momento continúa existiendo un río de papeleo de constatación. Los impulsos enviados mediante cinta o por cable mantienen los saldos bancarios del país, pero los cheques y valores siguen circulando como pruebas y para los estados de cuenta individuales. Parece como si fuese a disminuir el papel negociable, y con ello los robos de nómina y atracos a mensajeros de bancos; se difundirán delitos más sutiles, delitos que manipularán lo invisible y lo pasajero. Lo más frecuente serán las faenas hechas con ayuda de alguien de la casa, como ocurre en el mundo del papel y la tinta. Los computadores realizan tareas más voluminosas; sus atacantes son, con mayor frecuencia, grupos o equipos. Hasta ahora han sido principalmente jóvenes aficionados, gente conocedora de las técnicas informáticas, fortificados en general por la excusa de pedir prestado por un tiempo o de perjudicar sólo a alguna entidad corporativa impersonal.

Naturalmente, nos enteramos sólo de las martingalas que fracasan. El cálculo de las pérdidas anda en los Estados Unidos por los 300 millones de dólares anuales, extrapolados de manera verosímil, pero no firme, de menos de 300 casos. Se supone que el índice delictivo es un robo por año por cada 2000 computadores en servicio y cerca de medio millón de dólares por caso. Pero, ¿quién puede asegurarlo? Durante años, los auditores de cuentas en papel han conocido el fraude del redondeo hacia abajo, una treta para manipular centavos fraccionales sacándolos de muchas cuentas, sin que los totales muestren defecto porque todo está saldado. Se trata sólo de que el redondeo aleatorio hacia arriba o hacia abajo se ve reemplazado por el determinismo. Numerosas cuentas que se redondean hacia abajo se llevan honestamente, pero las demás las acumula preferencialmente el perpetrador. "¿Cuántos programadores se han retirado a una vida de ocio conforme sus programas, aceptados confiadamente y olvidados hace largo tiempo, continúan bombeando los centavos en sus cuentas a una velocidad cercana a la de la luz?". No incurriendo en la codicia indecente, esta paciente tenia de la moneda digital puede permanecer invisible.

Sea lo que sea aquello que yace en las profundidades, el picacho visible del iceberg está repleto de ingeniosos bribones, cuyas actividades no se deben tanto a la falsificación del soporte lógico (software) o del soporte físico (hardware) subversivo, sino a su talento personal como hombres de confianza. (En este elenco entran también algunas mujeres.) Veamos por ejemplo el juvenil jefe de Creative Systems Enterprises. Su pequeña compañía vendió unos cuantos centenares de miles de dólares de los mejores equipos de Western Electric. Obtuvo los suministros con reventa libre (todos fueron estampados "Reventa autorizada" en su propio taller) mediante un prolongado estudio de los procedimientos de las compañías telefónicas; hizo sus pedidos mediante Touchtone en correcto impreso de la compañía y recibió las entregas en el patio de equipos de la compañía en su furgoneta de la compañía, comprada en subasta de la compañía. Fue preciso pensar mucho, pero no ser gran experto en computadores. Después de que hubo penado lo suyo en cárcel, este sabihondo, por supuesto, llegó a ser asesor sobre seguridad en informática. Leal a su esti-

lo, podía "meterse en cualquier sistema de tiempo compartido del país".

He ahí en imagen y palabra el ingeniero de Silicon Valley que proyectó y cableó un computador especializado dentro de un cinto para llevar dinero, que lo ayudaba a contar las cartas durante las tallas de veintiuna real ("black jack") (1000 palabras de memoria de sólo lectura). Los conmutadores de entrada eran cuatro discos de desconexión rápida accionados por los dedos gordos de sus pies; la pantalla la formaban diminutos diodos emisores de luz incorporados en la parte interior del marco de sus gafas. Estaban demasiado cercanos como para que su ojo formara una imagen, pero podía distinguir los colores de los desenfocados "globos luminosos". Sin capital suficiente y menos suerte, perdió su apuesta, sufrió una conversión religiosa y dejó de jugar.

El mayor de los fraudes relatados fue una amplia conspiración de ejecutivos de los Angeles, que se daban la gran vida, quienes generaron falsos datos de entrada. Sacaron de la quiebra a la compañía de seguros en que trabajaban cargando sus archivos de disco con 64.000 pólizas falsas. Cuando sus complacientes auditores, un tanto ingenuos en una edad informática, pedían documentación de muestra, una oficina especial se ponía a trabajar a fin de crear los papeles necesarios para la inspección de la mañana siguiente. El dinero de verdad afluyó a través de ventas de acciones. El gran reventón, proveniente del soplo de un empleado despedido, terminó con 22 fallos condenatorios y un sinfín de pleitos, siendo de 2 millones de dólares la bolsa que detentaban accionistas y acreedores.

El más sutil y peligroso de los abusos informáticos no ha aparecido en aplicaciones de lucro ilícito. Fue más bien obra de la pura bravata dentro de un servicio universitario de tiempo compartido. Al seguirle la pista a una avería difícil de localizar, un programador despierto halló por casualidad una extraña secuencia de instrucciones en una impresión de programa de 400 páginas. La extraña lista permitía a "un terminal de las facultades... ser el gobernante definitivo e inequívoco del sistema". El virus lógico estaba escondido dentro de un programa de cintas editoras, ofrecido generosamente para uso público. Al cabo de medio año el programa editor fue pedido por un usuario que detentaba el nivel más elevado de privilegio dentro del sistema. La inserción, que había estado esperando esa oportunidad, se hizo cargo entonces, se colocó a sí misma dentro del centro del sistema, borró sus propias huellas y devolvió el control al

inocuo editor de cintas. El resultado sería completo sólo mucho más tarde, en un momento preordenado de gozoso triunfo, cuando el sistema dedicase todos sus terminales a imprimir 100 veces la magnífica inserción, luego el nombre del graduado inventor, a lo que seguiría "la peor avería en la historia del sistema".

Quizás el cuento sea mítico; no se ofrecen muchas pruebas. Los proyectistas de sistemas lo temen, sin embargo. La compleja arquitectura de ciertos sistemas nuevos incluye un ordenamiento jerárquico de capacidades. Ningún usuario puede emplear un programa a un nivel más alto que aquel que tal programa está destinado a alcanzar, aunque el usuario tenga los máximos privilegios. Sólo ciertas compuertas permiten un cambio de nivel aun cuando esté permitido. Las contraseñas son dobles y se cambian frecuentemente; a cada usuario se le saluda con un mensaje que dice el lugar y la fecha en que firmó por última vez. Pueden añadirse rutinas especiales, flexibles y cambiantes. Los sistemas militares incluso hacen sonar una alarma en la consola si fracasa un intento de acceso al computador.

El espionaje mediante líneas intervenidas es, por cierto, un problema constante, y un diseño normalizado del National Bureau of Standards ha tomado como base un ingeniosísimo sistema criptográfico, fundado en el trabajo de Claude E. Shannon, de hace ya una generación. Dicho diseño difunde ampliamente el mensaje a través de un rosario de símbolos. Hoy en día puede realizarse mediante unas cuantas pastillas –y muy en secreto, a menos que alguien haya cambiado subrepticiamente las pastillas. Los dos libros examinan asimismo los daños físicos infligidos a los ordenadores. Ni hablar de los pequeños imanes tradicionales para sabotear las cintas, sistema difícil de manipular. Si uno quiere fastidiar un computador, es difícil superar el método del incendio provocado. Con mayor satisfacción, uno puede emular a los protestatarios de Melbourne quienes "dispararon contra un computador con escopeta, lo que le ocasionó daños terribles. Fue una pérdida total". Un ciudadano sudafricano disparó cuatro veces a través de una ventana contra un computador de la recaudación de impuestos en Johannesburgo. Aunque magullado, el ordenador continuó trabajando para Hacienda.

En cierto modo, el sistema parece llevar la delantera, teniendo de su lado los batallones más nutridos. Lo que queda es un regusto de la duda de Bertolt Brecht: "¿Qué crimen es mayor: robar un banco, o fundar uno?" (P.M.).

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

ARMAS DE RADIACION INTENSIVA

THE EFFECTS OF NUCLEAR WEAPONS. Dirigido por Samuel Glasstone y preparado por el United States Department of Defense. Government Printing Office, 1977.

NUCLEAR POWER ISSUES AND CHOICES. Nuclear Energy Policy Study Group. Ballinger Publishing Company, 1977.

CUANDO EL MAR NEGRO SE DESECABA

HISTORY OF THE MEDITERRANEAN SALINITY CRISIS. Kenneth J. Hsü, Lucien Montadert, Daniel Bernoulli, Maria Bianca Cita, Robert E. Garrison, Robert B. Kidd y Frederic Mèlierés, Carla Müller y Ramil Wright en *Nature,* vol. 267, n.º 5610, págs. 399-403; 2 de junio de 1977.

INITIAL REPORTS OF THE DEEP SEA DRILLING PROJECT: Vol. 42, B. David A. Ross y Yu. Neprochnov. Government Printing Office, en prensa.

LA RADIACION COSMICA DE FONDO Y EL NUEVO MOVIMIENTO DEL ETER

THE FIRST THREE MINUTES: A MODERN VIEW OF THE ORIGIN OF THE UNIVERSE. Steven Weinberg. Basic Books, Inc., 1977.

DETECTION OF ANISOTROPY IN THE COSMIC BLACKBODY RADIATION. George F. Smoot; Marc V. Gorenstein y Richard A. Muller en *Physical Review Letters,* vol. 39, págs. 898-901; 1977.

RADIOMETER SYSTEM. Marc V. Gorenstein, Richard A. Muller, George F. Smoot y J. Anthony Tyson en *Review of Scientific Instruments,* vol. 49, n.º 4, págs. 440-448; abril, 1978.

EL NUCLEOSOMA

CELL AND MOLECULAR BIOLOGY. An appreciation. Eugene Ronseberg. Holt Rinehart & Winston, Inc., N.Y. 1971.

REFLECTIONS ON BIOCHEMISTRY. A. Kornberg, B.L. Horecker, L. Cornudella y J. Oró. Pergamon Press, 1976.

DIFFERENTIAL NUCLEASE ACTION ON NUCLEI & CHROMATIN FROM DEVELOPING GERM CELLS OF THE ECHINODERM *Holothuria tubulosa.* E. Rocha y L. Cornudella en *Biochem. Biophys. Res. Commun.,* vol. 68, n.º 4, págs. 1073-1081; febrero, 1976.

CAMBIOS EN LA ESPERMATOGENESIS Y DISTRIBUCION DE RIBOSOMAS EN EL EQUINODERMO *Holothuria tubulosa.* Luis Cornudella en *Avances de la Bioquímica,* eds. L. Cornudella, C.F. de Heredia, J. Oró y A. Sols. Salvat Editores, S.A., págs. 39-50, 1977.

EL SISTEMA DE APAREAMIENTO DE LEK DEL GALLO DE LAS ARTEMISAS

THE STRUT DISPLAY OF MALE SAGE GROUSE: A "FIXED" ACTION PATTERN. R. Haven Wiley en *Behaviour,* vol. 47, n.º 1 y 2, págs. 129-152; 1973.

EVOLUTION OF SOCIAL ORGANIZATION AND LIFE-HISTORY PATTERNS AMONG GROUSE. R. Haven Wiley en *Quarterly Review of Biology,* vol. 49, n.º 3, págs. 201-227; septiembre, 1974.

EL PRINCIPIO DE ADYACENCIA EN LA PERCEPCION VISUAL

VECTOR ANALYSIS IN VISUAL PERCEPTION OF ROLLING MOTION: A QUANTITATIVE APPROACH. Gunnar Johansson en *Psychologische Forschung,* vol. 36, n.º 4, págs. 311-319; 1974.

LOCAL AUTONOMY IN VISUAL SPACE. Walter C. Gogel y D.H. Mershon en *Scandinavian Journal of Psychology,* vol. 18, págs. 237-250; 1977.

UNIONES INTERCELULARES

CELL COMMUNICATION. Dirigido por R. P. Cox. John Wiley & Sons, Inc., 1974.

STRUCTURE AND FUNCTION OF INTERCELLULAR JUNCTIONS. L. Andrew Staehelin en *International Review of Cytology,* vol. 39, págs. 191-283; 1974.

FUNCTIONAL SIGNIFICANCE OF THE VARIATIONS IN THE GEOMETRICAL ORGANIZATION OF TIGHT JUNCTION NETWORKS. Barbara E. Hull y L. Andrew Staehelin en *The Journal of Cell Biology,* vol. 68, n.º 3, págs. 688-704; marzo, 1976.

GAP JUNCTION STRUCTURES, I: CORRELATED ELECTRON MICROSCOPY AND X-RAY DIFFRACTION. D. L. D. Caspar, D. A. Goodenough, Lee Makowski y W. C. Phillips en *The Journal of Cell Biology,* vol. 74 n.º 2, págs. 607-628; agosto, 1977.

TECNOLOGIA HIDRAULICA ROMANA

POWER. R. J. Forbes en *Studies in Ancient Technology: Vol. 2.* E. J. Brill, Leiden, 1955.

ROMAN DAMS IN TRIPOLITANIA. C. Vita-Finzi en *Antiquity,* vol. 35, págs. 14-20; 1961.

THE AQUEDUCTS OF ANCIENT ROME. Thomas Ashby, editado por I. A. Richmond. Clarendon Press, 1971.

MAN AND WATER: A HISTORY OF HYDRO-TECHNOLOGY. Norman Smith. Charles Scribner's Sons, 1975.

JUEGOS MATEMATICOS

EXPONENTIAL NUMBERS. Eric T. Bell en *The American Mathematical Monthly,* vol. 41, n.º 7, págs. 411-419; agosto-septiembre, 1934.

SOLUTION TO PROBLEM E 461. H. W. Becker en *The American Mathematical Monthly,* vol. 48, n.º 10, págs. 701-703; diciembre, 1941.

THE NUMBER OF PARTITIONS OF A SET. Gian-Carlo Rota en *The American Mathematical Monthly,* vol. 71, n.º 5, págs. 498-504; mayo, 1964.

THE THEORY OF PARTITIONS. George E. Andrews en *Encyclopedia of Mathematics and Its Applications: Vol. 2,* dirigido por Gian-Carlo Rota. Addison-Wesley Publishing Company, Inc., Advanced Book Program, 1976.

TALLER Y LABORATORIO

NATURAL AERODYNAMICS, R. S. Scorer. Pergamon Press, 1958.

GENERAL METEOROLOGY. Horace R. Byers. McGraw-Hill Book Company, 1959.

THE BEHAVIOUR OF CHIMNEY PLUMES. R. S. Scorer en *International Journal of Air Pollution,* vol. 1, n.º 3, págs. 198-220; enero, 1959.

SOME RESTRICTIVE METEOROLOGICAL CONDITIONS TO BE CONSIDERED IN THE DESING OF STACKS. Eugene W. Bierly y E. Wendell Hewson en *Journal of Applied Meteorology,* vol. 1, n.º 3, págs. 383-390; septiembre, 1962.